DISCUSSION PAPER SERIES

DP17400
(v. 5)

STARTUP TYPES AND MACROECONOMIC PERFORMANCE IN EUROPE

Ralph De Haas, Vincent Sterk and Neeltje van Horen

FINANCIAL ECONOMICS, INDUSTRIAL ORGANIZATION, INTERNATIONAL MACROECONOMICS AND FINANCE, MACROECONOMICS AND GROWTH AND PUBLIC ECONOMICS

CEPR

ISSN 0265-8003

STARTUP TYPES AND MACROECONOMIC PERFORMANCE IN EUROPE

Ralph De Haas, Vincent Sterk and Neeltje van Horen

Discussion Paper DP17400
First Published 21 June 2022
This Revision 25 October 2026

Centre for Economic Policy Research
187 boulevard Saint-Germain, 75007 Paris, France
2 Coldbath Square, London EC1R 5HL
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Financial Economics
- Industrial Organization
- International Macroeconomics and Finance
- Macroeconomics and Growth
- Public Economics

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Ralph De Haas, Vincent Sterk and Neeltje van Horen

STARTUP TYPES AND MACROECONOMIC PERFORMANCE IN EUROPE

Abstract

We construct a new dataset of 1.3 million European startups. Cluster analysis identifies five startup types present across countries, industries, and cohorts. Initial differences across types persist, and each displays a characteristic life cycle in productivity, employment, and survival. We embed these startup types in a firm dynamics model with heterogeneous entry, investment, and financing, and quantify how policy can shift cohort composition. Differentiated corporate tax reforms favoring high-performance types yield substantial productivity gains. Uniform fiscal instruments differ sharply: capital subsidies raise output and productivity through composition effects, labor subsidies reduce them, and interest subsidies are roughly neutral.

JEL Classification: D22, E24, E62, G32, G38, H25, L25, L26

Keywords: Startups, Firm dynamics, Capital structure, Corporate tax, Subsidies

Ralph De Haas - dehaasr@ebrd.com
EBRD, KU Leuven and CEPR

Vincent Sterk - v.sterk@ucl.ac.uk
Department of Economics, University College London and CEPR

Neeltje van Horen - n.vanhoren@uva.nl
University of Amsterdam and CEPR

Acknowledgements
We thank Filippo di Mauro for his enthusiastic support of this project, Victoria Marino for excellent research assistance, Marco Matani and Tommaso Bighelli for their expert help with running our code as part of constructing the 8th vintage of the CompNet dataset, and Tommaso Bighelli, Cagatay Bircan, John Haltiwanger, Sebnem Kalemli-Ozcan, Pete Klenow, Javier Miranda (discussant), Kjetil Storesletten and seminar participants at University College London, Institute for International Economic Studies (Stockholm), University of Regensburg, Danmarks Nationalbank, University of Bristol, University of Essex, Hong Kong University of Science and Technology, National University of Singapore, Hebrew University of Jerusalem, Toulouse School of Economics, Padova Macro Talks, Sao Paulo School of Economics, Nordic Summer Symposium in Macroeconomics, European Economic Association Annual Congress (Milan), and the CompNet ProdTalks for useful comments. The views expressed in this paper are ours and not necessarily those of the European Bank for Reconstruction and Development (EBRD).

Startup Types and Macroeconomic Performance in Europe*

Ralph De Haas†, Vincent Sterk‡, and Neeltje Van Horen§

October 23, 2026

Abstract

We construct a new dataset of 1.3 million European startups and cluster them on observable entry characteristics. Five types recur across countries, industries, and cohorts, each with a distinct life cycle in productivity, employment, and survival. A calibrated firm dynamics model with heterogeneous entrants and financial frictions quantifies how policy can improve macroeconomic performance by shifting startup composition. Budget-neutral corporate tax changes, differentiated across startup types, can yield substantial productivity gains. Subsidies applied uniformly across types differ sharply in their composition effects and aggregate implications: capital subsidies raise output and productivity, labor subsidies lower both, and interest subsidies are roughly neutral.

JEL classification: D22, G32, G38, H25, L25, L26, M13

Keywords: Startup heterogeneity, firm entry, capital structure, corporate taxation, productivity

*We thank Filippo di Mauro for his enthusiastic support of this project, Victoria Marino for excellent research assistance, Marco Matani and Tommaso Bighelli for their help with running our code as part of constructing the 8th vintage of the CompNet dataset, and Tommaso Bighelli, Çağatay Bircan, John Haltiwanger, Şebnem Kalemli-Özcan, Pete Klenow, Javier Miranda (discussant), Kjetil Storesletten and seminar participants at University College London, Institute for International Economic Studies (Stockholm), University of Regensburg, Danmarks Nationalbank, University of Bristol, University of Essex, Hong Kong University of Science and Technology, National University of Singapore, Hebrew University of Jerusalem, Toulouse School of Economics, Padova Macro Talks, Sao Paulo School of Economics, Nordic Summer Symposium in Macroeconomics, European Economic Association Annual Congress (Milan), and the CompNet ProdTalks for useful comments. The views expressed in this paper are ours and not necessarily those of the European Bank for Reconstruction and Development (EBRD).

†EBRD, KU Leuven, and CEPR. Email: dehaasr@ebrd.com.

‡University College London, NHH, and CEPR. Email: v.sterk@ucl.ac.uk.

§University of Amsterdam and CEPR. Email: n.vanhoren@uva.nl.

1 Introduction

Productivity growth remains anemic across many advanced economies (Akcigit and Ates, 2021, 2023; OECD, 2026b), prompting policymakers to seek novel levers to improve macroeconomic outcomes. A long-standing literature has explored several directions, including tax measures to stimulate R&D and innovation (Bloom, Griffith, and Van Reenen, 2002; Akcigit, Hanley, and Stantcheva, 2022) and structural reforms to reduce distortions in financial markets (Buera and Shin, 2017; Larrain and Stumpner, 2017), labor markets (Hopenhayn and Rogerson, 1993), and product markets (Edmond, Midrigan, and Xu, 2023). Such policies aim to raise the productivity of firms already in operation or to strengthen the selection process that weeds out inefficient ones (Syverson, 2011). This paper investigates a different and largely unexplored lever: influencing the types of new firms that enter the economy.

The idea of improving the composition of new firm cohorts—as opposed to "fixing" established generations—appears attractive for two reasons. First, startups are key drivers of job creation and productivity growth (Foster, Haltiwanger, and Krizan, 2001; Haltiwanger, Jarmin, and Miranda, 2013; Adelino, Ma, and Robinson, 2017).[1] Recent evidence (discussed below) also suggests that ex ante heterogeneity among new firms helps predict their later performance. Structural policies that shift the types of firms entering the economy may therefore generate substantial macroeconomic gains. Second, firm entry and exit rates are high, typically around 10 percent annually. Most of the firms that will operate twenty years from now do not yet exist, while many of today's firms will have disappeared.[2]

In practice, some governments already condition tax treatment or financial support for startups on observable characteristics such as firm size or financing structure, for instance through reduced corporate tax rates below certain size thresholds (Section 5.1.1). Yet size is

[1] We define startups broadly as all new firms that enter the economy, rather than narrowly as the subset of innovative firms that aim to scale up rapidly.

[2] For instance, the U.S. Longitudinal Business Database (LBD) shows that, in 2023, 73 percent of all firms were under 20 years old. The startup rate in that year was 9.1 percent. In Internet Appendix A, we show that European startup rates are comparable. Dent, Karahan, Pugsley, and Şahin (2016) use the LBD to show that about half of all sectoral employment reallocation in the United States since the end of the 1980s reflects changes in startup employment (the entry margin) rather than employment shifts among incumbent firms.

only one of several dimensions along which entrants differ, and not obviously the one most tied to later productivity. Firms that enter at a similar scale can look very different in other respects, and there is little evidence on which of those characteristics policy should condition on. Can policies improve macroeconomic performance by altering the mix of startup types, and if so, which types should be encouraged? Neither question has a clear answer.

We address these questions using large-scale administrative datasets for ten European countries: Croatia, Denmark, Finland, France, Italy, Lithuania, the Netherlands, Slovenia, Spain, and Sweden. The data contain a rich set of variables, derived from the balance sheets and income statements of individual startups. We collected them in close collaboration with the Competitiveness Research Network (CompNet), which uses a distributed microdata approach to produce regularly updated and internationally harmonized statistics on European firms. The result is a cross-country panel of more than 1.3 million startups, with representative coverage of the full startup population.[3]

A key challenge is that startup types are latent. We address this with k-means clustering, an unsupervised machine learning algorithm increasingly used in applied economics to handle unobserved heterogeneity. We take a startup's type to be revealed by the operating and financial configuration it starts out with, measured by employment, total assets, capital intensity, the cash-to-assets ratio, and leverage. These are the state variables of the firm dynamics model we develop below, and they are variables tax authorities already collect, so a policymaker could condition on them. The algorithm groups startups into five types—*Large*, *Capital-intensive*, *High-leverage*, *Cash-intensive*, and *Basic*—that capture much of the variation in these variables. A meta-clustering analysis shows that each type has similar characteristics across countries, and Monte Carlo simulations confirm that this similarity is not mechanical. The types recur across countries, sectors, and cohorts, and initial differences persist over a decade.

[3]This addresses limitations of other sources, including the under-representation of small firms in Compustat, the single-cohort design of the Kauffman Firm Survey, the limited variable coverage of the US Longitudinal Business Database, and the poor capture of firm entry and exit in Orbis (Bajgar, Berlingieri, Calligaris, Criscuolo, and Timmis, 2020; Kalemli-Özcan, Sørensen, Villegas-Sanchez, Volosovych, and Yeşiltaş, 2024).

We then document heterogeneity in performance along dimensions that do not enter the classification. Labor productivity and TFP differ markedly across types, and the gaps narrow but persist as cohorts age. The *High-leverage* type performs consistently poorly: even within the same country and sector, these firms are substantially less productive on both measures. The *Capital-intensive*, *Large*, and *Cash-intensive* types, by contrast, are typically more productive. Shifts in the composition of startup cohorts may therefore carry sizable macroeconomic consequences.

To interpret these patterns and quantify their policy implications, we calibrate an equilibrium firm dynamics model in which types differ in technology and initial equity. Entry, financing, and hiring decisions are jointly determined. With type-specific parameters calibrated to target each type's life-cycle profiles in the clustering variables and in profits, the model rationalizes the performance differences as the joint outcome of technological heterogeneity (TFP, returns to scale, fixed costs) and differences in initial equity. Policy changes then affect aggregate outcomes not only through the responses of incumbents but also through shifts in the composition of entering cohorts. This composition effect is governed by three empirically estimable objects: the life-cycle profile of each type, the type-specific entry elasticities, and the immediate effect of the policy on firm profits. We recover each from the data, so the counterfactual results can be traced back to their empirical inputs.

We conduct two sets of counterfactual policy experiments. First, we quantify the "policy space" opened by a budget-neutral differentiation of corporate income tax rates across startup types. Such a reform alters the incentives of different types to enter and thus shifts the composition of new cohorts.[4] A budget-neutral reform that changes corporate tax rates by no more than ten percentage points at the type level can raise aggregate labor productivity by up to 4.5 percent when entry elasticities are calibrated at the one-digit sector level and by up to 1.2 percent at the two-digit level. The largest gains come from raising the tax on *Basic*

[4]Earlier work has shown that corporate taxation affects firm entry (Da Rin, Di Giacomo, and Sembenelli, 2011) and the size of new firms (Da Rin, Di Giacomo, and Sembenelli, 2010), and that tax policy more broadly shapes the funding available to young firms (Edwards and Todtenhaupt, 2020).

startups, which finances lower taxes on other types. These estimates abstract from re-sorting: because types are defined by entry choices, type-contingent differentiation could lead firms to change the configuration they enter with, a margin the model holds fixed.

Second, we use the model to compare three uniform, budget-neutral subsidies financed through corporate-tax adjustments: a capital subsidy, a labor subsidy, and an interest subsidy for young firms. Although none of these instruments targets a particular type, each still changes the mix of entrants, because the types differ in capital intensity, labor share, and reliance on debt. A capital subsidy raises aggregate output and labor productivity despite the fall in entry caused by the financing tax. An equally costly labor subsidy reduces both, while a much cheaper interest subsidy is close to neutral in the aggregate, though it tilts composition toward *High-leverage* startups.

Both sets of experiments show that the macroeconomic effects of fiscal policy on new firms depend in large part on its interaction with startup heterogeneity. This holds not only for instruments that explicitly condition on type, but also for uniform ones, whose effects operate through differences in exposure that representative-firm frameworks do not capture.

Related literature. We build on an emerging literature that documents the importance of ex ante heterogeneity for firms' performance over their life cycle. This heterogeneity manifests itself in vastly different growth expectations among new entrepreneurs (Hurst and Pugsley, 2011), predictability of firm performance from observable characteristics of founders and firms at entry (Schoar, 2010; Guzman and Stern, 2015; Belenzon, Chatterji, and Daley, 2017; Guzman and Li, 2023), strong cohort effects over the business cycle (Sedláček and Sterk, 2017), and systematic differences in financial structure at entry that reflect firm and founder characteristics (Robb and Robinson, 2014). Quantitatively, ex ante heterogeneity accounts for most variation in firm-level employment, as shown by Sterk ⓡ Sedláček ⓡ Pugsley (2021).

We contribute to this literature in four ways. First, we ask how policy can use ex ante firm heterogeneity to improve macroeconomic outcomes. This shifts the focus from the

widely studied "startup deficit"—a decline in firm entry observed in several countries over the past decades—to the *composition* of new startup cohorts as a determinant of aggregate productivity and employment.[5] This links the composition of entrant cohorts to the micro origins of productivity dispersion (Hsieh and Klenow, 2009; Restuccia and Rogerson, 2013; Hsieh and Klenow, 2014; Midrigan and Xu, 2014; Moll, 2014; Gopinath, Kalemli-Özcan, Karabarbounis, and Villegas-Sanchez, 2017; Brandt, Kambourov, and Storesletten, 2026; De Loecker and Syverson, 2021).

Second, we treat ex ante heterogeneity as a multidimensional object. Most existing studies focus on a single dimension, such as firm size.[6] We instead jointly characterize firms by five key variables at startup. Two of these, leverage and cash holdings, describe the initial financial conditions of firms. Earlier work documents the importance of financing choices at startup (Robb and Robinson, 2014; Derrien, Mésonnier, and Vuillemey, 2025) and the persistence of leverage differences across firms (Lemmon, Roberts, and Zender, 2008). We combine these financial variables with real ones such as scale and capital intensity. Clustering on all five therefore reveals distinct combinations of financial and operating characteristics within the entering population.

Our classification also uses only information observable in a firm's first year of operation, unlike the latent types recovered in Albert and Caggese (2021) and Sterk ⓡ Sedláček ⓡ Pugsley (2021). Relative to classifications based on founders' growth intentions (Hurst and Pugsley, 2011), innovation orientation (Botelho, Fehder, and Hochberg, 2026), or founding strategy (Guzman and Stern, 2015; Guzman and Li, 2023), ours is coarser in economic content. Those distinctions nonetheless show up in our types: *Basic* startups (half of all entrants) are the counterpart of Hurst and Pugsley's (2011) undifferentiated small businesses, while the calibrated fixed costs and initial equity of *Large* and *Capital-intensive* startups match the

[5] On the startup deficit, see Decker, Haltiwanger, Jarmin, and Miranda (2017); Alon, Berger, Dent, and Pugsley (2018); Karahan, Pugsley, and Şahin (2024); Biondi, Inferrera, Mertens, and Miranda (2023).

[6] Bernard, Dhyne, Magerman, Manova, and Moxnes (2022), using Belgian data on firm-to-firm trade, emphasize the role of production-network characteristics in explaining firm-level heterogeneity. Other work examines ex ante differences across founders, including cognitive and non-cognitive personality traits (Levine and Rubinstein, 2017) and age (Azoulay, Jones, Kim, and Miranda, 2020).

profile of Schoar's (2010) transformational entrepreneurs.

Third, we obtain empirical estimates for the elasticity of firm entry and show that these elasticities are heterogeneous across startup types. Such estimates are rare, even though they are a standard input in firm dynamics models, which often assume either fixed entry (that is, a zero elasticity) or a free entry condition (that is, an infinite entry elasticity).[7] Our estimates can thus be used to impose more empirical discipline on models in the tradition of Hopenhayn (1992) and models with firm entry more generally.

Fourth, we embed empirically identified, ex ante heterogeneous startup types in a firm dynamics model. This allows us to compare the macroeconomic effects of differentiated tax reforms as well as uniform fiscal instruments within a single framework.[8] The model captures compositional channels from which representative-firm frameworks abstract. These channels turn out to be quantitatively first-order and can reverse the sign of the effects such frameworks predict. A capital subsidy, for instance, raises aggregate output, even though the fall in entry it causes would on its own reduce output.

2 Data and clustering

2.1 Startups in the CompNet database

We use confidential microdata collected in close collaboration with the Competitiveness Research Network (CompNet), which maintains an annually updated, internationally harmonized dataset for European countries. To preserve firm-level confidentiality and improve comparability, CompNet uses the distributed microdata approach of Bartelsman, Haltiwanger,

[7] Other estimates can be found in Sedláček and Sterk (2017) and Gutiérrez, Jones, and Philippon (2021). Yet, they rely on full-blown structural DSGE models to estimate the entry elasticity, whereas our method is agnostic. Moreover, we estimate elasticities for different startup types.

[8] Our differentiated-tax experiment also connects to the literature on size-dependent policy, which shows that policy thresholds tied to observable firm characteristics can distort the firm-size distribution (Guner, Ventura, and Xu, 2008; Garicano, Lelarge, and Van Reenen, 2016; Ewens, Xiao, and Xu, 2024). Where that literature measures the cost of size-dependence, we study how conditioning on the entry balance sheet can shift the composition of new startup cohorts and subsequent macro outcomes.

and Scarpetta (2004): standardized code is sent to national statistical agencies, which run it on their confidential firm data and return key moments that describe the distribution of firm characteristics by sector, year, and country. Particular care is taken to ensure cross-country consistency, enabling international comparisons.[9]

To collect data on startups, we embed our own code in the standard instructions sent to national authorities for the eighth CompNet vintage. This code runs within each national environment and identifies every firm that first becomes economically active in a given country and year. The year of first activity is the start year and a cohort consists of all firms with the same start year. We also observe each firm's formal registration year and drop observations if the lag between registration and first activity exceeds four years, or if registration occurs *after* the start year (a handful of cases). We exclude firms that enter with more than 50 employees: such entrants are unlikely to reflect genuine firm creation and instead capture spin-offs or subsidiaries of existing firms. This yields a cross-country panel of 1,345,489 startups established during 2000–2019.[10]

To preserve confidentiality, CompNet does not permit the extraction of individual firm observations and instead releases statistics for cells containing at least ten firms. For each country–industry–type–cohort, we track variables by firm age, so the unit of observation is a country–industry–type–cohort–age cell. Industries follow either a one-digit or two-digit NACE Rev.2 classification.[11] For each cell, we compute the mean of total assets, employment, capital intensity, cash ratio, leverage, wages, labor productivity, TFP, and profit margin,

[9]We exclude Germany, Poland, and Slovakia, whose national sources omit firms below 20 employees (10 in Poland). Coverage is unbalanced across countries (Appendix Table A1), but identification comes from within country–industry–cohort comparisons. Lopez-Garcia and di Mauro (2015) provide more details on CompNet.

[10]We drop firms not observed every year between startup and exit. CompNet is not restricted by legal form and targets the population of firms with at least one employee, subject to country-specific reporting thresholds. Because our clustering variables require complete balance-sheet information, the effective sample consists mostly of incorporated, limited-liability firms, which are the entities required to file full accounts.

[11]The one-digit industries are administration, construction, hospitality, ICT, manufacturing, professionals, trade, and transport. While the two-digit partition exploits finer within-sector variation, the one-digit partition retains more of the population, since CompNet drops cells with fewer than ten firms and this roughly halves the two-digit sample. We lead with the one-digit classification because our questions concern aggregate outcomes and because it preserves representativeness, but report one- and two-digit results throughout.

as well as the exit rate.[12] This cell-level aggregation matches our empirical specifications (lifecycle and entry-elasticity regressions) in which high-dimensional fixed effects absorb systematic cross-country, cross-industry, and cross-cohort differences.

Internet Appendix A shows that our firm-creation counts are broadly consistent with Eurostat; Biondi et al. (2023) find the same for total firm populations. Entry sizes match external evidence too: averaged across entrants, startup employment is in the low single digits, consistent with evidence that firms are very small at birth. Haltiwanger, Jarmin, and Miranda (2013) and Moreira (2017), for example, report that U.S. entrants average around four employees. Appendix Figure A2 shows that CompNet entry employment closely tracks Eurostat benchmarks in nearly all countries. Appendix Table A2 describes all variables.

2.2 Identifying startup types

We categorize firms within each country by k-means cluster analysis (Everitt, Landau, Leese, and Stahl, 2011), a form of unsupervised machine learning to approximate latent heterogeneity by a finite set of groups (Bonhomme, Lamadon, and Manresa, 2022). We cluster startups on five variables measured in the first year of operation: the number of employees, total assets, capital intensity (fixed assets per employee), cash to total assets, and leverage (total debt to total assets).[13] Together these variables characterize the operating and financial configuration with which a firm starts.[14] We view this configuration not as a primitive of the entrepreneur's problem (nobody sets out to found "a capital-intensive firm") but as an

[12]Labor productivity is real value added per employee. TFP is based on a homogeneous production function estimated at the two-digit sectoral level via the control function of Ackerberg, Caves, and Frazer (2015), aggregated using employment shares.

[13]Earlier work has assessed the role of startup scale as measured by assets or employees (Albuquerque and Hopenhayn, 2004; Kerr and Nanda, 2010; Buera, Kaboski, and Shin, 2011) or set-up costs (Derrien, Mésonnier, and Vuillemey, 2025); liquidity and cash holdings (Bolton, Wang, and Yang, 2019); and leverage (Robb and Robinson, 2014; Derrien, Mésonnier, and Vuillemey, 2025). We also use capital intensity as a clustering variable because heterogeneity in production functions is expressed in different capital-to-labor ratios (Oberfield and Raval, 2021).

[14]The five variables are not independent: capital intensity is fixed assets per employee, linking it to employment and, through fixed assets, to total assets. We retain all five variables because each isolates a distinct margin (labor-market scale, production technology, balance-sheet size, financing structure, and liquidity). We convert each variable to a z-score so that units do not drive the clustering.

equilibrium object shaped by the founder's technology and ambition, the sector, and the financing on offer.[15] This is sufficient for our purposes, provided the configuration a firm chooses at entry is informative about its latent type.

We choose these variables for three reasons. First, they are the state variables of the firm dynamics model in Section 4. Clustering on them describes entry heterogeneity in the model's own terms and lets us recover the structural parameters governing each type: TFP, returns to scale, fixed costs, and initial equity. Second, they are the margins fiscal instruments act on: capital allowances apply to fixed assets, interest deductibility to debt. Types defined on these variables therefore differ in their exposure to the instruments we evaluate. Third, they are recorded in standard financial statements, so a policymaker can observe them at entry and for the whole entrant population.

The clustering itself is standard. Let x_i collect the clustering variables for firm i, expressed as z-scores. The algorithm partitions firms into k clusters $C_1, \ldots, C_k$ to solve $\min_{\{C_j\}} \sum_{j=1}^{k} \sum_{i \in C_j} \|x_i - \bar{x}_j\|^2$, where $\bar{x}_j$ is the mean of cluster C_j. To gauge the number of clusters we compute $\eta^2 \equiv 1 - WSS/TSS$, where $WSS \equiv \sum_{j=1}^{k} \sum_{i \in C_j} \|x_i - \bar{x}_j\|^2$ and $TSS \equiv \sum_{i=1}^{n} \|x_i - \bar{x}\|^2$, with $\bar{x}$ the unconditional mean.[16]

Letting k run from 1 to 10, η^2 reaches about 0.6 at $k = 5$, so five clusters capture roughly three-fifths of the variation in the clustering variables. Beyond that, η^2 continues to rise, but only slowly. Because within-cluster dispersion falls mechanically as k grows, neither η^2 nor the scree plots (Appendix Figure A5) unequivocally pin down an optimal k. We settle on five because the partition is stable under refinement: raising k to six or seven leaves four of the five clusters largely intact and mainly subdivides the residual fifth, which splits by scale (Section 3.1 and Appendix Table A3).

None of this means that startup heterogeneity is discrete. Following Bonhomme, Lamadon, and Manresa (2022), we treat the five types as an approximation of continuous variation. This

[15] For example, Ma, Murfin, and Pratt (2022) show that financial constraints affect young firms' capital acquisition, with startups relying disproportionately on used physical capital.

[16] This statistic has a similar interpretation to the R^2 reported in regression analysis.

approximation is valid when the heterogeneity is low-dimensional and the sorting variables are informative about it (conditions we assess in Section 3). The partition sorts entrants along the margins fiscal policy acts on, which is what the counterfactuals require.

Two caveats remain. First, we do not observe innovative capacity or founder human capital. Our types therefore capture a firm's underlying quality only to the extent that it shows up in the first-year accounts. Since tax authorities cannot observe founder quality either, this limitation matters less for the policy question we study. Second, the entry-year balance sheet is measured after registration and may reflect both the founder's intended configuration and frictions encountered during entry. Entry conditions can leave a lasting mark on later scale (Moreira, 2017). Frictions of this kind, however, are unlikely to reproduce the same five configurations across countries, sectors, and cohorts.

2.3 Comparing startup clusters across countries

We run the cluster analysis separately for each country, so neither the resulting clusters nor their shares need be similar across countries. To assess whether they are, we run a second-stage "meta-clustering" analysis, which groups comparable clusters from different countries. This also provides an objective procedure to assign common names to similar clusters. Specifically, we repeat the clustering analysis while taking the cluster centers derived from each country's first-stage clustering as the observations.[17]

The four panels in Figure 1 summarize this meta-clustering. Different meta-clusters are shown in different colors. The panels show five clusters arising in each country. We label them *Basic*, *Large*, *Capital-intensive*, *Cash-intensive*, and *High-leverage*, and characterize each in Section 3.1. For example, the first panel contains a red meta-cluster of *Capital-intensive* startup clusters. Each of the individual red circles indicates a country-level cluster of startups that stand out nationally because of their high capital intensity. Such a distinct

[17]That is, in the meta-clustering procedure, the units of observation are the z-scores of the first-stage cluster centers, averaged across years and industries. The z-scores are computed within each country to allow for measurement differences and institutional variation across countries.

Capital-intensive cluster emerges in each country and the algorithm bunches them together in a single meta-cluster. The meta-clustering yields $\eta^2 = 0.96$: country-level clusters of the same type differ far less from one another than clusters of different types do within a country.

We next check whether this similarity could arise mechanically, using a Monte Carlo experiment in which we replicate our procedure on artificial data in which no clusters exist. Concretely, for each country we consider 1,000 i.i.d. draws from a log-normal distribution for the cluster variables, with means and standard deviations as observed in the data. We then apply exactly the same two-step procedure as to the actual data: clustering within each country, followed by meta-clustering across countries. Each time we compute η^2. Appendix Figure A6 shows that these η^2 statistics are much lower in the experiment than the 0.96 observed in the real-world data. The Monte Carlo experiment hence supports our interpretation that the cluster outcomes in the data, and their similarity across countries, capture economic correlation structures rather than being mechanical artifacts of the clustering procedure.

3 An anatomy of startup types

3.1 Five startup types

Table 1 summarizes our clustering analysis and reports the five types together with their population shares: *Basic* (49 percent), *Large* (4 percent), *Capital-intensive* (7 percent), *Cash-intensive* (26 percent), and *High-leverage* (14 percent). These labels reflect the key dimension along which a startup type differentiates itself. For example, *Large* startups employ, on average, 20 people when they begin operations, compared with only two to four for the other types. Likewise, *Cash-intensive* startups hold, on average, 54 percent of their assets as cash when they commence operations, whereas this is just 12 percent for other startups. Startup types consistently display their defining characteristic across countries, as the first panels of Appendix Figure A7 illustrate.

Raising k above five leaves the classification largely intact. Appendix Table A3 reports

cluster means for $k = 6$ and $k = 7$ alongside the baseline $k = 5$: the *Large*, *Capital-intensive*, *Cash-intensive*, and *High-leverage* types retain their defining characteristic at every k, with mean capital intensity of the *Capital-intensive* type above 90,000 euros per employee, the cash ratio of the *Cash-intensive* type above 0.5, the leverage ratio of the *High-leverage* type above 1.1, and the *Large* type retaining by far the largest balance sheet. The additional clusters mainly split the residual *Basic* type by scale rather than reorganize the full classification. At $k = 6$, *Basic* separates into a group with about three employees and a group with about 18 employees but relatively few assets. At $k = 7$ a further intermediate group emerges.

Figure 2 (Panel A) shows that while the clustering algorithm yields the same startup types in each country, their local prevalence differs somewhat. For example, *Cash-intensive* startups are relatively common in Italy but less so in France. This variation is informative: it shows that the algorithm concentrates *Cash-intensive* entrants precisely in the country where firms are independently documented to hold high cash balances, especially in early life.[18] Because the same five types emerge everywhere, such institutional differences shift the *composition* of entrants across a fixed set of types, rather than the definitions of the types themselves.[19] Panels B and C show that the same types are also present across industries and cohorts, although their prevalence varies. The classification thus captures recurring dimensions of startup heterogeneity rather than a country-, sector-, or cohort-specific pattern.

We next examine the distribution of startup types across industries in more detail. Appendix Figure A8 shows that all five types are present across one-digit industries in individual countries, while Appendix Figure A9 shows that they are also present in nearly all two-digit industries. We then classify all two-digit sectors into one of nine groups. We

[18]Italian company law requires subscribed capital to be paid in cash at incorporation: at least a quarter of the minimum for an ordinary limited liability company (S.r.l.), and the full amount for the single-member and simplified forms. New firms therefore begin with cash on the balance sheet before deploying it into fixed assets. Italian firms also hold relatively high precautionary cash buffers, partly reflecting financial constraints and fragmented local banking (Dottori and Micucci, 2018; Fasano and Deloof, 2021).

[19]As noted in Section 2.1, the set of contributing countries changes across cohorts. Changes in aggregate shares before 2009 and after 2017 therefore partly reflects countries entering or leaving the sample rather than shifts in underlying type shares. Over 2009–2017 the same nine countries contribute every cohort, and type shares over this period are stable.

follow the Eurostat classification for the technology intensity (manufacturing) or knowledge intensity (services) of industries, as proxied by their R&D expenditure/value added. Appendix Figure A10 (Panel A) shows that the five startup types are consistently present across sectors with different technology and knowledge intensities. In Panel B, we follow Mian and Sufi (2014) and assign two-digit industries to one of four categories: construction, tradable, non-tradable, and other. The five startup types are again present across these economic sectors. *Capital-intensive* startups are somewhat more common among tradable firms, and *High-leverage* ones among non-tradable firms. These patterns also clarify what our classification adds relative to industry codes alone. If startup types merely relabeled sectors, each would concentrate in a few industries. This is clearly not the case. Startup types therefore capture differences between firms that sector classifications largely miss. For the same reason, policies that differentiate only by sector would do little to shift the composition margin we study.

A related question is what unites the firms within each cluster at a deeper economic level, given that they span various industries. As shown above, types do not proxy for industry. Indeed, the same four sectors (retail trade, wholesale trade, food and beverage services, and specialized construction) rank among the five largest industries in each type, and no two-digit industry accounts for more than 20 percent of any type. The model in Section 4 instead attributes the differences between types to production technology (scalability and TFP) and initial equity. For example, *Capital-intensive* startups in retail trade and specialized construction hold roughly ten times as much fixed capital per employee as *Basic* startups in the same industries, which the calibrated model traces to a technology that is highly scalable in capital but not in labor.

3.2 Early life cycles: Persistent differences

In Figure 3, we explore how startups develop as they age, using a regression approach to estimate life-cycle profiles in the cluster variables.[20] We find that type differences narrow over time but remain large and statistically significant. First, *Cash-intensive* startups begin with substantially more cash (relative to total assets) than other types and draw it down over time, but even after 12 years their cash intensity remains about 10 percentage points higher. Second, *High-leverage* startups are initially about 40 percentage points more leveraged than *Basic* ones (even within the same country and industry), a gap that shrinks to 5 points after 12 years. Third, *Capital-intensive* startups begin with about 47,000 euros more fixed assets per employee than *Basic* ones, a gap that narrows to about 18,000 euros after 12 years. Fourth, by contrast, *Large* startups diverge: their employment lead over *Basic* startups widens from 18 to 24 employees, while their capital intensity converges.

Convergence could in principle reflect regression to the mean rather than genuine economic adjustment. Several features of our setting make this explanation unlikely. First, classification depends on a combination of five variables, making it difficult for measurement error in any single variable to generate the type-specific profiles we observe. Second, not all variables converge: employment for *Large* startups diverges over time. Pure regression to the mean would shrink gaps in all dimensions. Third, as we show next, startup types also differ persistently along performance variables *not* used in the clustering, which regression to the mean in the clustering variables cannot explain.

3.3 Performance of the five startup types

Table 2 uses panel regressions to explore performance differences across startup types. Observations refer to cell averages for all new firms in a country, one-digit industry, startup

[20] We regress each cluster variable on dummies for four startup types (using the *Basic* type as benchmark), country × cohort fixed effects, and two-digit industry fixed effects, using the full panel dataset at the two-digit industry level. We run a separate regression for each age group (one-year-old firms, two-year-old firms, etc.) and plot the successive coefficients for the startup type dummies. Appendix Figure A11 shows that the life cycle patterns are very similar when aggregating data up to the one-digit industry level.

year (cohort), age, and type. For example, an observation could refer to the average productivity of Spanish ICT firms of the *High-leverage* type, established in 2005, at age seven. We include dummies for the startup types, using *Basic* firms as the omitted group. High-dimensional fixed effects (FE) flexibly control for unobservable drivers of firm performance that might correlate with startup type. Country × cohort FE absorb time-invariant characteristics common to startups established in a specific country and year; industry × cohort FE control for time-invariant traits common to startups in a specific industry and established in a given year; and country × industry FE absorb all time-invariant variation of startups in a specific industry and country. Finally, we add interactions between startups' age and their country, industry, or cohort. This captures startup traits that are specific to the country, industry, or cohort and that vary with firm age.[21]

The results in Panel A of Table 2 show that *Cash-intensive* and *Large* firms outperform *Basic* ones on labor productivity (column 1), TFP (column 2), likelihood of early exit (column 3), and wages (column 4). The same holds for *Capital-intensive* firms, which display the highest labor productivity of all types. *Large* startups stand out for low exit rates but also lower profit margins (column 5), while *Cash-intensive* firms have relatively high TFP and profit margins. *High-leverage* firms, by contrast, consistently and strongly underperform on labor productivity and TFP, and operate with lower profit margins even though they pay lower wages than other startup types. In Panel B of Table 2, we replicate this analysis using only firms between five and nine years old. The results line up closely with those in Panel A, indicating that the performance divergence between startup types is not solely driven by transient differences in early life.

Lastly, in unreported cross-sectional linear-probability regressions, we examine which startups achieve "gazelle" status by their fifth year.[22] We classify gazelles as startups

[21] We provide robust but unclustered standard errors as we use data on the full population of startups across industries and countries rather than a random sample of these (Abadie, Athey, Imbens, and Wooldridge, 2023). All results are robust to clustering at the industry-country-cohort level.

[22] These regressions include country × cohort FE, industry × cohort FE, and country × industry FE and can only be run for Italy and Spain as they require extracting additional information from CompNet.

that demonstrate above-median average TFP growth during their first five years while simultaneously reaching the top quintile of the overall TFP distribution by age five. Consistent with our discussion on the importance of initial firm size, only *Large* startups demonstrate an above-average probability of attaining gazelle status.

4 Startup composition: A theoretical framework

4.1 Model environment

We now present a firm dynamics model calibrated to the empirical patterns documented above. The model serves two purposes. First, by comparing type-specific parameter values, it gives a structural interpretation of the five types (Section 4.3). Second, it enables the policy counterfactuals in Section 5, where we evaluate budget-neutral differentiation of corporate tax rates across types alongside three uniform fiscal instruments (capital, labor, and interest subsidies). Because the incidence of these instruments varies across types, they shift cohort composition, with markedly different macroeconomic consequences.

The model period is one year and we denote firm age by a. We distinguish between three age categories: entrants, young firms, and old firms. The age at which firms turn old is denoted a^*. Each entrant belongs to one of a finite number of types. Startups decide on an initial capital investment that cannot be reversed in subsequent years, in the spirit of the literature on lumpy investment (Doms and Dunne, 1998; Cooper and Haltiwanger, 2006). Entrants also face a financial friction, which shapes how they set up the financial structure of the firm. Old firms are not financially constrained.[23] Firms maximize the present value of profits, with a discount factor $\Lambda \in (0, 1)$.

Technology and real rigidities. Firms use capital, k, and labor, l. Their production function is given by $y = A_a k^\alpha l^\gamma$, where A_a denotes TFP, varying by age and startup type.

[23]This assumption eases computations but is by no means crucial (especially since our focus is on young firms). Also, Figure 4 shows that after a decade, financial leverage is low among all types.

The output elasticities $\alpha > 0$ and $\gamma > 0$ vary across types but not age. Capital depreciates at a rate $\delta \in (0, 1)$ which is common across firms. Production requires a fixed cost $c \geq 0$, which may differ across types. The wage w is common across firms. Entrants decide on an initial level of capital but cannot further adjust the capital stock until they turn old at age a^*, after which they can adjust freely. We abstract from productivity shocks.

Assets and financial constraints. Firms can invest in short-term liquid assets (cash), denoted m, which earn an interest rate $r = \frac{1}{\Lambda} - 1$. However, entrants cannot borrow in this type of asset: they face a constraint $m_0 \geq 0$. This constraint does not apply to old firms, which are therefore financially unconstrained. Entrants are born with an initial amount of equity m_{-1}, which may vary across types.

To finance the initial capital investment, entrants can take out a long-term loan $b \geq 0$ with a duration a^*. The loan is repaid in equal installments at a per-period interest rate $r_b(b)$. We assume that $r_b(b)$ is weakly increasing in the loan amount b, that $r_b(b) \geq \widetilde{r}$, and that $r_b(0) = \widetilde{r}$. Here, $\widetilde{r} = \frac{1}{\sum_{a=0}^{a^*-1} \Lambda^{a+1}} - \frac{1}{a^*}$ is the frictionless rate for this repayment schedule, consistent with the subjective discount factor and the interest rate on cash. In other words, whenever the firm borrows, it pays an external premium that increases in loan size. This captures, in simplified form, the external finance premium in models of financial frictions, such as costly state verification (e.g., Townsend, 1979).

Old firms ($a \geq a^*$). Old firms are free to adjust the capital stock and they face no financial constraints. For simplicity, we assume that once a firm reaches old age, TFP remains constant and firms exit at an exogenous rate ρ. Because old firms face no constraints, their decision problem is equivalent to one in which they statically rent capital and labor from households. No arbitrage implies that the rental rate of capital is given by $r + \delta$.[24] The profits of an old firm are given by $\pi = Ak^\alpha l^\gamma - (r + \delta) k - wl - c$. Old firms statically choose k and l to maximize profits. The firm value (excluding assets) is given by $V^* = \frac{\pi}{1-\Lambda(1-\rho)}$. Old firms are

[24] Equivalently, we could assume that the firms own the capital.

indifferent between holding assets (cash and/or capital) or paying dividends.

Young firms ($a < a^*$). Young firms cannot adjust their long-term loan and capital stock, but they must fulfill loan obligations and pay for the upkeep of the capital stock. Young firms do not pay any dividends, i.e. they retain all profits in the form of cash.[25] For simplicity, we assume that firms do not exit during this phase. The young firms' decision problem can be expressed as:

$$\begin{aligned} V_a(m_{a-1}, k, b) = \max_{l_a} \Lambda V_{a+1}(m_a, k, b) \\ s.t.\ m_a = (1+r)m_{a-1} + \pi_a - \left(\frac{1}{a^*} + r_b(b)\right) b, \\ \pi_a = A_a k^\alpha l_a^\gamma - \delta k - w l_a - c, \\ l_a \geq 0. \end{aligned}$$

The value of a firm turning old, including assets, is given by $V_{a^*}(m_{a-1}, k, 0) = V^* + k + (1 + r)m_{a^*-1}$. Note that at that point, the firm has fully repaid its loan ($b = 0$).

Entrants ($a = 0$). Just after entry, a firm chooses the amount of capital, long-term debt, and cash to maximize firm value subject to a budget constraint and financial constraints:

$$\begin{aligned} V(m_{-1}) = \max_{k,b,m_0} V_0(m_0, k, b) \\ s.t.\ k + m_0 = m_{-1} + b, \\ b, m_0 \geq 0, \end{aligned}$$

where the initial amount of cash, m_{-1}, is exogenous.

[25] If the firm is currently constrained, or may be constrained in the future, it is optimal not to pay dividends. It may be the case that a young firm is fully unconstrained. It is then indifferent between paying dividends and accumulating cash. For those cases, we assume that the firm does not pay dividends.

Startup types and entry decisions. We index firm types by $j = 1, 2, .., J$. Types differ in technology $(\{A_a\}_{a=0}^{a=a^*}, \alpha, \gamma, c)$ and initial equity (m_{-1}). From now on, we denote by $V_j \equiv V(m_{-1}; \{A_a\}_{a=0}^{a=a^*}, \alpha, \gamma, c)$ the entry value of a firm of type j.

For any type j, there are a certain number of *potential* entrants in any given period. Potential entrants know their type and, conditional on type, the expected firm value V_j. However, they do not know any idiosyncratic realization that may distinguish them from other entrants of the same type. The only source of within-type heterogeneity that matters for the entry decision is the entry cost $\theta_{i,j}$, which varies across potential entrants. Both the number of potential entrants and the distribution of the entry costs may vary across types.

Optimal decision-making implies that a firm of type j starts whenever $V_j \geq m_{-1} + \theta_{i,j}$. That is, for each type there exists a cutoff value of $\theta_{i,j}$ such that only potential entrants with an entry cost below the cutoff start a firm.[26] We can now express the actual number of entrants of type j as $n_j = g(V_j, \Psi_j)$ where g is an increasing function in V_j, and where Ψ_j denotes the number of potential entrants of type j and their distribution over entry costs. Entry therefore depends on expected value and shocks to the potential-entrant pool.

Households and equilibrium. Adding a household sector to the model allows us to solve for the general equilibrium. We focus on steady-state (i.e., long-run) equilibria. The household has a subjective discount factor Λ, so that in the steady state $r = \frac{1}{\Lambda} - 1$ is given as above. We assume for tractability that households have Greenwood–Hercowitz–Huffman preferences over leisure and consumption. Optimal labor supply then takes the form $l = \kappa_0 w^{\kappa_1}$, where κ_1 is the labor supply elasticity, which can be calibrated externally, and κ_0 is a scaling parameter. Labor market clearing requires: $l = \sum_j n_j \left(\sum_{a=0}^{a^*-1} l_{a,j} + \frac{l_{*,j}}{1-\rho} \right)$ while the asset and capital market clear trivially.

[26] *Within* a startup type, all entrants have the same expectations prior to drawing the demand and production function. This implies there are no entry selection effects within types (only across types). While this modeling assumption is standard in the literature following Hopenhayn (1992), we verify it empirically in Section 4.4.

4.2 Calibration

4.2.1 Post-entry parameters

Recall that startup types differ in terms of technology $(\{A_a\}_{a=0}^{a=a^*}, \alpha, \gamma, c)$ and initial equity (m_{-1}). Our goal is to interpret the five types by finding a parametrization that fits the empirical profile of each type. We can then compare the parameter values for different types to understand how their fundamentals differ. The quantitative solution method for the model is discussed in Internet Appendix B.

For each type, we feed into the model the empirically observed TFP profiles by age, plotted in the top left panel of Figure 4. We externally set $\Lambda = 0.96$, corresponding to an annual real interest rate of four percent, and $\delta = 0.1$, a depreciation rate of ten percent. We further assume that $r_b(b) = \widetilde{r} + \zeta b$, where we calibrate ζ to target a 10 basis point spread based on average debt across all categories in the data. We set the wage, and scale the levels of TFP, such that the model matches (by construction) the capital-to-labor ratio and the employment level of old *Basic* firms.[27] Next, we calibrate for each type the parameters $(m_{-1}, \alpha, \gamma, c)$, targeting the respective empirical profiles, presented in Figure 4. That is, we target the profiles for employment, capital intensity, leverage, the cash ratio, and profits. The latter we include because profits enter directly into the entry condition, which is at the center of the policy counterfactuals. We further set $a^* = 10$ and $\rho = 0.03$.

4.2.2 Entry elasticities

To study the fundamentals driving the post-entry outcomes for each type, we only need to calibrate the parameters described above. However, in order to conduct policy counterfactuals, we also need to set pre-entry parameters that determine how the number and type composition of entrants reacts to a change in firm values. To this end, a first-order approximation (in logs) of this function gives the following expression for the percentage change in the number

[27] For this initial calibration step, we assume $\alpha = 0.3$ and $\gamma = 0.4$, close to the values found in the next step. Also, the parameter value for the interest rate function is $\zeta = 2.27 \times 10^{-6}$.

of entrants of type j:

$$d \ln n_j = \varepsilon_j \cdot d \ln V_j + \psi_j \tag{1}$$

Here, $\varepsilon_j > 0$ is the elasticity of the number of entrants of type j with respect to the ex ante expected firm value, which is given by the mass of entrants at the entry cutoff, relative to the mass of entrants below the cutoff. Moreover, ψ_j denotes shocks to the number and distribution of potential entrants of type j (that is, shocks to Ψ_j). Thus, the number of entrants of a certain type can rise either through higher expected value or through a shock to the potential-entrant pool.[28]

To estimate entry elasticities, we run a regression based on the entry condition, Equation (1). A challenge is that in the data we observe the ex post average realization, denoted $\ln \tilde{V}_j$, while it is the ex-ante expected value $\ln V_j$ that pins down the entry. The difference between the two arises due to common shocks that occur *after* entry. If we were to estimate Equation (1) based on ex post realizations, the residual would contain $\ln \tilde{V}_j - \ln V_j$, giving rise to correlation between the residual and the right-hand side variable, $\ln V_j$. This would introduce a bias. A second issue is that, under a null hypothesis of free entry (a baseline assumption in the theoretical literature), the entry elasticity is infinite, making the regression specification ill-defined. To circumvent these problems, we rearrange Equation (1) as:

$$d \ln V_j = \frac{1}{\varepsilon_j} d \ln n_j + \xi_j + u_j \tag{2}$$

where $\xi_j \equiv -\frac{\psi_j}{\varepsilon_j}$ is a fixed effect due to shocks, for which we control in the regression. Moreover, $u_j \equiv d \ln \tilde{V}_j - d \ln V_j$ is the gap between realized and expected value, which arises from shocks realized only after entry. Rearranging in this way moves that gap into the residual rather than the regressor (we return to the resulting identifying assumptions below).

We use observations at the country(c)-industry(i)-type(j)-cohort(t) level, with correspond-

[28] An alternative setup to arrive at the same result is to assume an unlimited number of entrants and to assume that the entry cost is homogeneous across entrants, but an increasing function of the number of entrants (within a type). In that case, the elasticity ε_j is determined by the shape of the entry-cost function. See, for example, Gutiérrez, Jones, and Philippon (2021).

ing indices between brackets.[29] We control for shocks to the number and distribution of entrants at the country-industry-type level and at the country-industry-cohort level by using interactive fixed effects at these levels. The former is important as the number of potential entrants may naturally vary by country, industry, and type. The latter captures that, in a specific industry and country, the number of potential entrants may fluctuate over time. We run these regressions both at the one-digit and at the two-digit industry level. This results in the following specification, estimated via ordinary least squares (OLS):

$$\ln V_{c,i,j,t} = \beta_{0,j} + \beta_{1,j} \ln n_{c,i,j,t} + \xi_{c,i,j} + \xi_{c,i,t} + u_{c,i,j,t} \tag{3}$$

We note that $\frac{1}{\beta_{1,j}}$ is the type-specific entry elasticity and that, under free entry (that is, an infinite elasticity), it holds that $\beta_{1,j} = 0$. We estimate Equation (3) as a single pooled regression across the five startup types, interacting the intercept and the slope with type indicators. The slope coefficients $\beta_{1,j}$ are therefore type-specific, while the fixed effects $\xi_{c,i,t}$ are common across types, so that they absorb shocks shared by all types within a country-industry-cohort. We measure $\ln V_j$ using each startup type's life-cycle profile for profits and exit rates, together with the number of entrants for each country-industry-type-cohort in our dataset.[30] We assume a discount factor $\Lambda = 0.96$, corresponding to an annual discount rate of about 4 percent. We use only cohorts observed for at least seven years (that is, those established before 2011). For age twelve and onward, we assume that profits and the year-on-year exit rate remain fixed.[31] Figure 5 reports the estimated inverse entry elasticities by startup type, using both one- and two-digit industry classifications.

[29] Both the regressand (log firm value) and the regressor (log number of entrants) are defined at the cell level, so the regression involves no inference across levels of aggregation.

[30] We reconstruct profits as the product of the EBIT margin (EBIT per unit of revenue) and revenue, at the country-industry-type-cohort level.

[31] That is, we compute firm value at age zero as: $V_0 = \sum_{a=0}^{11} \Lambda^a s_{0,a} \pi_a + s_{0,12} \pi_{12} \sum_{a=12}^{\infty} \Lambda^a s_{11,12}^{a-12} = \sum_{a=0}^{11} \Lambda^a s_{0,a} \pi_a + \Lambda^{12} s_{0,12} \frac{\pi_{12}}{1 - \Lambda s_{11,12}}$, where $s_{k,l}$ is the survival rate between age k and $l \geq k$. We verified that changing the age cutoff at which profits are held constant hardly affects the results. This reflects that the component of the firm value beyond age twelve is relatively unimportant for the startup value, due to discounting and high exit rates at young ages.

Because the value measure is based on each type's survival profile, and survival is measured relative to the cohort's entrant count, the constructed $\ln V_{c,i,j,t}$ and the regressor $\ln n_{c,i,j,t}$ share a component: a cohort's entrant count. This mechanical link could in principle bias the estimated slopes. To check that this does not drive the cross-type differences on which the composition experiments of Section 5 rely, we recompute firm value using survival profiles held constant across cohorts within each country and startup type, so that a cohort's own entrant count no longer enters its value measure. The estimated entry elasticities remain close to the baseline, and their ordering across startup types is unchanged.

Identification. The slope in (3) recovers the inverse entry elasticity under two assumptions. First, the fixed effects absorb shocks to the pool of potential entrants: $\xi_{c,i,t}$ removes pool and profitability shocks common to all types within a country-industry-cohort, and $\xi_{c,i,j}$ removes permanent differences in the size and cost of each type's pool of potential entrants. Second, conditional on these, the residual variation in the number of entrants is orthogonal to the post-entry surprises in $u_{c,i,j,t}$. This holds by timing: entry is set on value expected at founding, whereas $u_{c,i,j,t}$ contains shocks realized only afterwards. One qualification is that the fixed effects do not absorb type-specific, time-varying shocks to the pool of potential entrants. Such shocks raise entry without raising expected value and would bias $\beta_{1,j}$ downward. They would reorder the cross-type elasticities only if their importance relative to value shocks differed systematically across types.

4.3 Interpreting the five startup types

We now use the calibrated model to interpret the startup types.[32] We present the model fit in Figure 4 and parameter values in Table 3. In the table, we divide initial assets and the fixed cost by the wage per worker for ease of interpretation. The model accounts reasonably well for the empirical profiles, considering that for each startup type we target 50 moments

[32]For this exercise, the estimated entry elasticities are irrelevant.

with just four parameters.[33] As for the parameter values, we observe large differences in startup types both in terms of their physical structure (production function, fixed cost, TFP) and their startup equity. This observation, combined with the satisfactory model fit, suggests that heterogeneity in technologies and initial equity are major drivers of differences between startup types. We now consider in more detail how the fundamentals of the five types differ.

Basic startups. *Basic* firms have moderate TFP and moderate returns to scale ($\alpha+\gamma = 0.7$). As they age, their TFP and employment levels increase, and their capital intensity (capital-to-labor ratio) declines. *Basic* types in the model finance their initial investment entirely from equity and so carry no debt, whereas in the data their leverage is low (0.23 at entry) but not zero. Their cash holdings are modest in both. The presence of many *Basic* startups, with an undifferentiated strategy and mediocre performance, fits with recent work on the productivity gap between, on the one hand, a small group of well-performing firms that operate at the technological frontier and, on the other hand, a long tail of less-productive laggards (Hsieh and Klenow, 2009; Andrews, Criscuolo, and Gal, 2016). Our clustering analysis indicates that such an uneven productivity distribution already emerges when new firm cohorts are born.

Large startups. *Large* firms have the highest overall scalability of all types ($\alpha + \gamma = 0.75$). Strictly speaking, we model scalability as a technological feature, but it could also be interpreted as scalability of customer demand (Hottman, Redding, and Weinstein, 2016). The TFP levels of *Large* startups are also relatively high, suggesting they combine physical and human capital in a relatively productive way. These fundamentals make entrants of this type poised to grow even bigger. Indeed, they have almost seven times as many employees as other startups, and their *level* of capital is highest among all types (their capital intensity is the second highest, after the *Capital-intensive* startups). In the data, we observe that *Large* firms are more prevalent in sectors like the manufacturing of basic materials (metals,

[33] One dimension along which the model fits the data less well is that it predicts a cash ratio profile for *Cash-intensive* that is slightly increasing, whereas in the data this profile is flat. Note further that the TFP profiles are the same in the data and the model, by construction.

paper). *Large* startups also remain persistently larger than the other types, consistent with empirical findings in Sterk ⓡ Sedláček ⓡ Pugsley (2021), who use US data to document the importance of ex ante heterogeneity for firm-level employment.

The scalable technology of *Large* firms comes with fixed costs more than four times as high as those of *Basic* startups. The high initial fixed investment and high fixed cost imply a strong financing need. Indeed, according to the model, the initial equity of *Large* entrants is almost six times as high as that of *Basic* firms. This substantial initial equity allows *Large* firms to operate without high leverage and with some cash holdings, in line with the data.

Capital-intensive startups. Compared to *Basic* startups, the scalability of *Capital-intensive* firms is much higher for capital (α) but lower for labor (γ). Capital intensity (the capital-to-labor ratio) is hence much higher. As they become more productive and hire more workers over time, the capital intensity of these firms remains relatively elevated.[34]

The fixed cost of *Capital-intensive* firms is almost three times that of *Basic* firms, consistent with their prevalence in manufacturing, transport, hospitality, and rental & leasing. These high capital requirements are reflected in initial equity almost three times that of *Basic* firms. This equity finances the initial capital stock, leaving *Capital-intensive* firms with relatively little cash, especially early in life (again in line with the data).

Cash-intensive startups. These startups are less scalable, especially with respect to capital ($\alpha = 0.24$). On the other hand, they have the highest TFP of all types, while their fixed cost is the lowest. These may be businesses producing "niche" goods or services, being small but profitable. While *Cash-intensive* firms are present across all sectors, they are more common in ICT, professional services, and construction.

According to the model, the initial equity of these firms is somewhat lower than that of *Basic* ones but larger than that of *High-leverage* startups. Their low scalability with respect

[34]In the data, the capital intensity of these startups is particularly high in the first year, which the model does not fully capture. This may reflect adjustment costs in hiring workers, which we do not include in the model.

to capital means that *Cash-intensive* startups initially invest relatively little. This allows them to retain a significant amount of initial equity as cash on their balance sheet and to avoid high leverage. Indeed, these firms initially hold more than half of their assets as cash, about five times as much as other startup types.

High-leverage startups. In terms of scalability, *High-leverage* startups resemble *Basic* ones, but they have significantly lower TFP, especially when young. According to the model, *High-leverage* firms enter with the lowest initial equity of all types, so they borrow substantially to finance their first investments: their average leverage ratio at entry is 118 percent in the data.[35] As they age, they pay off this debt and their leverage almost converges to that of other types after a decade. This deleveraging aligns with Dinlersoz, Kalemli-Özcan, Hyatt, and Penciakova (2019) and Kochen (2025), while its gradual nature is in line with Lemmon, Roberts, and Zender (2008), who document that leverage heterogeneity is persistent.

Relation to canonical typologies of entrepreneurship. While we impose no prior taxonomy, the estimated primitives align well with key startup classifications from the entrepreneurship literature. Hurst and Pugsley (2011) separate the mass of small businesses with no intention to grow from the minority that aims to scale. Schoar (2010) makes a similar distinction between subsistence and transformational entrepreneurship. In our classification, *Basic* startups (half of all entrants) correspond to the first group: they have moderate TFP, moderate returns to scale, and modest equity. In contrast, *Large* and *Capital-intensive* startups (jointly 11 percent of entrants) have the features that define transformational entry: high scalability (overall for *Large* startups, in capital for *Capital-intensive* ones), fixed costs

[35] A leverage ratio above one reflects negative book equity, which arises when first-year operating losses exceed the initial equity injection. The lower bound of zero on the equity injection m_{-1} constrains the *founding* balance sheet, whereas entry leverage in the data is measured from the first filed accounts at the end of the first operating year. Our model captures the underlying mechanism qualitatively: once cumulative operating losses drive a young firm's cash position negative ($m_a < 0$), the firm carries short-term debt and its measured leverage rises. *High-leverage* firms combine the lowest initial equity (Table 3) with the lowest early-life TFP, which is why the model generates the highest leverage for this type. The model nonetheless understates its level: initial leverage of *High-leverage* firms is well below one in the model, against values above one in the data (Figure 4).

roughly three to five times those of *Basic* entrants, and initial equity roughly three to six times as high. That this group is small, and recurs across countries and cohorts, matches the stylized fact that transformational entrepreneurship is rare.

Recovering this classification from the data, rather than imposing a taxonomy, is what lets us measure population shares, life-cycle profiles, and entry elasticities for each type (quantities that intention-based classifications cannot measure at population scale). It also partitions entrants into groups that respond differently to entrepreneurship policies, as our counterfactual experiments require.

4.4 Within-type selection at entry

A key assumption of the model is that entry selection operates across types, not within them. We now test this assumption empirically. We run age-partitioned regressions in which observations are at the country–two-digit sector–type–cohort level (for example, Spanish beverage manufacturers of the *Cash-intensive* type established in 2017). For each age, we regress a performance metric (labor productivity, TFP, exit rate, or employment) or a cluster variable (total assets, capital intensity, cash intensity, or leverage) on the log number of firms in the cohort, to estimate the relationship between cohort size and later startup performance. We saturate the specification with country-sector-type, country-sector-cohort, country-type-cohort, and sector-type-cohort fixed effects, and cluster standard errors at the country-sector-cohort level. If positive within-type selection were material, larger cohorts would systematically differ in later performance and composition. Appendix Figures A12–A13 show that the estimated elasticities are economically small, statistically insignificant at most horizons, and show no systematic trend, at both the two- and one-digit sector definitions. Within-type entry selection thus appears to be empirically negligible. This evidence supports holding the mapping from type to productivity fixed as entry responds to policy. Section 5.1 discusses what it does and does not imply for type-contingent taxation.

5 Counterfactual policy experiments

We now use the model to study how fiscal policy shifts the composition of new startup cohorts and, through it, aggregate macroeconomic outcomes. We run two sets of experiments. The first, a corporate income tax differentiated by startup type, traces the composition channel in its sharpest form. The second, three uniform subsidies (on capital, labor, and interest), does *not* condition on type, yet still shifts composition because each subsidy benefits types unevenly.

5.1 A budget-neutral tax policy differentiated by startup type

We start by introducing a corporate income tax. The after-tax profits of a firm are now given by $\pi_{a,j} = (1-\tau_{\pi,j})\left(A_a k_{a-1}^{\alpha} l_a^{\gamma} - k_a + (1-\delta)k_{a-1} - w l_a - c\right)$, where $\tau_{\pi,j}$ is the corporate tax rate which may vary across types. Our goal is to examine the aggregate effects of a differentiated corporate income tax rate. To impose discipline, we restrict ourselves to policies that are revenue-neutral.[36] This implies that, if some startup types are to be taxed less, other types need to be taxed more. We further impose equilibrium, i.e., the labor market clearing condition holds. Lastly, we limit ourselves to policies of no more than a 10-percentage-point change in the corporate tax rate for each type. Aside from these restrictions, we search the entire space of policy changes differentiated by type, and evaluate the macroeconomic implications for any admissible policy.

Our policy experiments are based on a numerical Monte Carlo procedure. We consider a large number of uniform random draws for the tax policies. For any draw, denoted $\tilde{\tau}_{\pi,j}$, we set the tax rates as $\tilde{\tau}_{\pi,j} + \bar{\tau}$, where $\bar{\tau}$ is a uniform adjustment factor used to ensure budget neutrality. We then jointly solve for firm decisions, $\bar{\tau}$ and the wage w such that the policy is budget neutral and the labor market clears.

Based on the estimated entry elasticities, we evaluate the change in the number of firms

[36] I.e., we impose that $\sum_a \sum_j n_{a,j} \frac{\tau_{\pi,j}}{1-\tau_{\pi,j}} \pi_{a,j} = 0$, where $n_{a,j}$ is the number of firms of age a and type j.

within each type. This allows us to construct counterfactuals for the number of startups by age and type, i.e. $n_{a,j}$. Similarly, we solve for the counterfactual choices in each of these bins. We then aggregate to compute counterfactual aggregate outcomes.

After evaluating the macro outcomes for the entire space of admissible policies, we group policies into bins according to their intensity, as measured by the maximum absolute change in the tax rate across types. The idea is that larger policy changes may bring about greater macroeconomic effects, but may also face stronger political resistance. It is therefore useful to study the policy space conditional on a certain intensity of the policy. For each bin, we compute policies that bring about the largest positive and the largest negative change in a macro variable (aggregate productivity or employment). We label this the policy frontier and present the tax-rate changes that generate the policy upper bound. We perform this analysis twice: once using entry elasticities derived from one-digit sector data and once using elasticities based on two-digit sector data (Figure 6). The two-digit approach exploits more granular variation in entry elasticities, whereas the one-digit approach retains more of the startup population, since two-digit cells with fewer than ten underlying firms are dropped under CompNet confidentiality rules.

Because our types are defined by entry choices (employment, capital intensity, total assets, leverage, and the cash ratio), a schedule that taxes types differently changes the incentive to enter as one type rather than another. Section 4.4 showed that within-type selection is empirically negligible, but that evidence comes from an environment in which taxes are not type-contingent and so cannot rule out re-sorting under a schedule that rewards type membership. If entrants could re-sort into a favorably taxed type without matching its fundamentals, part of the measured shift would be relabeling rather than genuine reallocation, and the realized gain would be smaller. We therefore interpret the differentiated-tax gains as a ceiling. The uniform-subsidy experiments of Section 5.2 do not condition on type and are hence immune to this concern.

Are large aggregate gains possible? Figure 6 summarizes our first, differentiated policy experiment. Panels A and B present the policy space for aggregate labor productivity (using entry elasticities based on a one-digit and two-digit sector classification, respectively).[37] The horizontal axis measures the intensity of the potential policy change (that is, the maximum change in the tax rate across all firms). Moving from left to right, warmer colors indicate stronger corporate tax-rate differentiation. The solid (dashed) line plots the policy upper bound (lower bound). This is the largest possible aggregate labor productivity increase (decrease) given a certain policy intensity.

The figure shows that substantial macroeconomic gains are possible. For the maximum policy intensity of 0.1 (that is, when the tax rate does not change by more than 10 percentage points for any type) labor productivity can increase by 4.5 percent under one-digit entry elasticities (Panel A) and 1.2 percent under the more granular two-digit elasticities (Panel B). Yet substantial macroeconomic losses are also possible for policies that shift the composition away from high-productivity startup types. The maximum productivity decline exceeds 4 percent under one-digit and 2 percent under two-digit elasticities (Panels A and B).

Panels C and D of Figure 6 again plot the policy space, this time for aggregate employment. Small employment gains are possible: 0.2 percent for a policy intensity of 0.1 (Panel C). These gains are absent when using the two-digit sector classification (Panel D). At the same time, larger employment losses—of 1.6 percent (Panel C) and 1 percent (Panel D)—are possible. This asymmetry reflects that the high-productivity types the optimal policy encourages are relatively capital-intensive, so they boost output per worker more than headcount, while taxing *Basic* startups (the type with the highest entry elasticity and the largest population share) can sharply reduce the total number of firms.

Which startup types to encourage? The two panels on the left of Figure 7 plot the corporate tax policies associated with the policy upper bounds (using the one-digit sector

[37]We focus on labor productivity rather than TFP, since the former also captures the macro effects of differences in capital investment between different types. That said, the exercise can also be performed for TFP (or any observable outcome). Figure 8 considers TFP alongside several other outcomes.

classification). Likewise, the two panels on the right show the associated firm shares. Which startup types should policy encourage to increase productivity or employment? The upper left panel shows that, to achieve the productivity upper bound, taxes should be lowered for *Capital-intensive*, *High-leverage*, and *Large* types, financed by a tax increase for *Basic* types, while the tax on *Cash-intensive* startups changes little. The upper right panel shows that, as a result, the share of *Capital-intensive* startups increases, as do the shares of *Cash-intensive*, *High-leverage* and *Large* startups. Accordingly, the share of *Basic* startups falls. Since labor productivity is relatively low among *Basic* firms, the shift in startup composition increases aggregate labor productivity.

The lower left panel of Figure 7 shows the tax rates associated with the employment frontier. To boost aggregate employment, taxes should be lowered for the *Basic* type, and increased for the *Cash-intensive*, *High-leverage*, and *Large* types. This leads to an increased share of *Basic* startups (lower right panel). Intuitively, the entry elasticity among *Basic* startups is relatively high. A tax cut for this type therefore has a relatively large impact on the overall number of entrants, and therefore employment.

Policy trade-offs. The previous analysis suggests a trade-off between maximizing productivity and employment. To gauge this more directly, Figure 8 provides scatter plots that depict pairs of changes in aggregate or average macroeconomic outcomes due to each potential policy.[38] The top panel shows there is some trade-off between increasing aggregate employment and labor productivity: the two outcomes are negatively correlated across policy changes. Yet, there exist specific reforms that simultaneously boost (or reduce) aggregate labor productivity and employment, e.g., increasing the shares of *Large* and *Cash-intensive* startups as much as possible (given a certain policy intensity). There is also a fairly strong negative trade-off between aggregate employment and TFP (middle panel). The trade-off between aggregate labor productivity and TFP is more limited (bottom panel).

[38] We again analyze policies with a maximum intensity of 0.1 while using entry elasticities based on the one-digit sector classification.

5.1.1 Tax differentiation for startups in practice

Our analysis shows that tax differentiation across startup types could yield substantial macroeconomic gains. Such differentiation should in principle be feasible, since it requires distinguishing startups on only a few observable traits, although the re-sorting concern discussed above could limit the gains in practice. That said, several countries have implemented policies that, either implicitly or explicitly, distinguish between startup types (or between startups and more mature firms). Belgium, France, Latvia, Lithuania, Luxembourg, the Netherlands, Poland, Portugal, Slovakia, and Spain all apply reduced corporate tax rates to small firms, typically defined by a profit or turnover threshold (OECD, 2026a).

A few countries differentiate tax rates among startups specifically. Spain, for example, taxes newly created companies at 15 percent in their first two profitable years, excludes new companies that belong to a corporate group, and extends the reduced rate to four years for certified innovative startups (PwC, 2026). In India, the rate applying to manufacturing startups depends on their annual turnover (Kalra, 2019). Others differentiate implicitly, through tax credits, breaks, and exemptions that substantially lower effective rates for young and small businesses. In some cases, fiscal support is also conditioned on how a startup is funded: credit guarantees and subsidized loans on the debt side, and tax relief for investors purchasing new shares on the equity side (such as the U.S. capital gains tax exclusion for shares in qualifying small companies).

Our results provide insight into how tax policies that penalize large firms can adversely affect startup composition and reduce macroeconomic performance, consistent with the size-dependent-policy literature, though here the cost operates through the entry margin rather than the scale of incumbents. Similarly, size-contingent regulations (such as less onerous labor legislation for small firms) may adversely affect the composition of new-firm cohorts, as they effectively discourage the entry of larger startups. Existing schemes that ease access to debt or equity matter through the same channel. We find that *High-leverage* startups perform substantially worse than other types, so policies that facilitate startups'

access to credit may affect firm performance quite differently from measures to help startups issue equity (see also Section 5.2).

5.2 Subsidies on labor, capital, and interest

We now consider three uniform fiscal instruments that do not condition on type: a subsidy per unit of installed capital, a labor subsidy that reduces the effective wage paid by firms, and an interest subsidy that lowers the borrowing rate for young firms (those below age ten). While uniformly applied, the instruments affect types differently and hence shape startup composition and macroeconomic outcomes. Internet Appendix B describes how we introduce these instruments into the model.

5.2.1 Capital subsidy

First, we consider a flow subsidy on capital, denoted $\tau_k < 0$, per unit of installed capital. To finance the subsidy, we assume that corporate income tax rates are adjusted uniformly, so that the overall intervention is budget neutral. The policy is scaled such that financing it requires a five-percentage-point increase in the corporate income tax rate, yielding a capital subsidy of 0.29 percent. As before, we impose that wages adjust to clear the labor market.[39]

Table 4 presents the effects of this subsidy. Panel B shows that it raises the share of *Capital-intensive* startups (the type that benefits most). Panel C turns to aggregate effects. The total number of firms declines, since financing the subsidy requires a higher corporate income tax. The decline is concentrated among *Basic* firms, whose entry elasticity is relatively high. The aggregate capital stock rises, and output and labor productivity also rise. Panel C also decomposes the macroeconomic effects into those coming from the number of startups, those from their composition, and a remainder. The first two are quantitatively significant, and they go in opposite directions: fewer firms reduce aggregate output and employment,

[39]Subsidies of this kind are widely used. Criscuolo, Martin, Overman, and Van Reenen (2019), for example, show that discretionary investment grants in the United Kingdom raised investment and employment among small firms, with no detectable effect on TFP.

while the shift in composition raises them. The remainder reflects the within-type response, as firms of each type install more capital when it is subsidized. This intensive margin accounts for most of the rise in the aggregate capital stock and offsets most of the decline in employment caused by the smaller number of firms. Net of the composition effect, however, output would fall slightly, so it is the composition channel that makes the output effect positive.

5.2.2 Labor subsidy

Next, we consider a subsidy on labor, denoted $\tau_l < 0$. The effective wage paid by the firm is $w(1+\tau_l)$, while households continue to receive w. As before, we impose labor market clearing and budget neutrality, and scale the subsidy such that it requires a five-percentage-point increase in the corporate income tax rate.

Table 4 shows that this subsidy has sizable negative effects on both output and labor productivity. The output decline is driven by the fall in the number of firms and a negative composition effect. Both are quantitatively large. The composition effect mainly reflects a fall of 1.5 percentage points in the share of *Capital-intensive* startups. The decline in labor productivity has different sources: the composition effect contributes about -1.4 percent, and the within-type response, as firms of each type hire more workers at the lower effective wage, contributes about -1.7 percent. Although the exercise considers a labor *subsidy*, the results suggest that labor *taxes* may be less detrimental to the macroeconomy than often assumed: if the effects are roughly symmetric, using higher labor taxation to lower corporate income taxes may raise output through changes in the number and composition of startups.

5.2.3 Interest subsidy

Finally, we consider an interest subsidy, denoted $\tau_b < 0$, for young firms (firms below age ten). We consider a subsidy of 0.29 percent, the same magnitude as in the capital subsidy experiment. The policy requires a small increase in the corporate income tax rate. As shown in Panel C of Table 4, the macroeconomic effects of the policy are small but positive. The

number of startups increases, which raises output. However, this increase is partly offset by a negative composition effect, as the policy increases the share of *High-leverage* startups, which exhibit relatively poor performance in both the data and the model.

6 Conclusions

This paper combines novel and comprehensive data on startups in multiple European countries with a firm dynamics model to quantify the macroeconomic gains from policies that shift the composition of new startup cohorts. Using unsupervised machine learning, we find that similar clusters of startups (identified by the choices they make at entry) emerge in each country. Moreover, the subsequent performance of these types (in terms of employment, productivity, and survival) differs markedly. There are therefore potential macroeconomic gains (or losses) to be had from policies that alter the composition of new startup cohorts.

Two sets of policy experiments based on the calibrated model quantify these impacts. First, a budget-neutral differentiation of corporate income tax rates (bounded at ten percentage points per type) can raise aggregate labor productivity by up to 4.5 percent using one-digit entry elasticities, and by up to 1.2 percent using two-digit elasticities. The *Basic* startup type has the highest entry elasticity, so taxing it more heavily produces the largest compositional shift toward the higher-productivity types. Second, even uniform fiscal instruments differ in their macroeconomic effects once compositional channels are accounted for. A capital subsidy raises output and labor productivity despite the fall in entry caused by the financing tax, while an equivalently sized labor subsidy reduces both. An interest subsidy aimed at young firms is roughly neutral in the aggregate but tilts composition toward *High-leverage* startups, which perform poorly in both data and model.

Our results speak directly to current policy debates on how to revive business dynamism. In the United States, concerns about secular declines in startup entry and the disappearance of high-growth young firms have prompted a combination of industrial policy and deregulatory

responses. Laws such as the CHIPS and Science Act and the Inflation Reduction Act provided subsidies and tax credits that reduce fixed costs and ease financing constraints for firms, including new entrants, in capital-intensive sectors. In Europe, similar ambitions are reflected in the EU Inc. regulation proposed by the European Commission in March 2026, which would create an optional, EU-wide corporate legal form, designed mainly with startups and scale-ups in mind, as an alternative to the 27 national regimes, while leaving corporate taxation to member states. Our framework provides a tool to assess the macroeconomic effects of such reforms, with particular emphasis on their impact on the composition of entering firms. More broadly, our approach connects to the renewed interest in industrial policy (Juhász, Lane, and Rodrik, 2024; Criscuolo et al., 2019): by explicitly modeling how policies shape the composition of startup cohorts, it enables a quantitative evaluation of both the benefits and potential distortions associated with targeted interventions.

Increasingly, statistical agencies make rich micro data on firms accessible to researchers, so our methodology can readily be applied to a wider set of countries. A natural question is to what extent differences in startup composition account for cross-country differences in macroeconomic performance. Richer startup classifications could also sharpen the policy implications. A related extension is to design optimal combinations of fiscal instruments, rather than evaluating them one at a time, and to examine how the optimal mix depends on a country's initial startup composition. We leave these issues for future research.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge. 2023. "When Should You Adjust Standard Errors for Clustering?." *Quarterly Journal of Economics* 138 (1): 1–35.

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. 2015. "Identification Properties of Recent Production Function Estimators." *Econometrica* 83 (6): 2411–2451.

Adelino, Manuel, Song Ma, and David Robinson. 2017. "Firm Age, Investment Opportunities, and Job Creation." *Journal of Finance* 72 (3): 999–1038.

Akcigit, Ufuk, and Sina T. Ates. 2021. "Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory." *American Economic Journal: Macroeconomics* 13 (1): 257–98.

Akcigit, Ufuk, and Sina T. Ates. 2023. "What Happened to US Business Dynamism?." *Journal of Political Economy* 131 (8): 2059–2124.

Akcigit, Ufuk, Douglas Hanley, and Stefanie Stantcheva. 2022. "Optimal Taxation and R&D Policies." *Econometrica* 90 (2): 645–684.

Albert, Christoph, and Andrea Caggese. 2021. "Cyclical Fluctuations, Financial Shocks, and the Entry of Fast-growing Entrepreneurial Startups." *Review of Financial Studies* 34 (5): 2508–2548.

Albuquerque, Rui, and Hugo A. Hopenhayn. 2004. "Optimal Lending Contracts and Firm Dynamics." *Review of Economic Studies* 71 (2): 285–315.

Alon, Titan, David Berger, Robert Dent, and Benjamin Pugsley. 2018. "Older and Slower: The Startup Deficit's Lasting Effects on Aggregate Productivity Growth." *Journal of Monetary Economics* 93: 68–85.

Andrews, Dan, Chiara Criscuolo, and Peter Gal. 2016. "The Best versus the Rest. The Global Productivity Slowdown, Divergence across Firms and the Role of Public Policy." Productivity Working Papers No. 5, Organisation for Economic Co-operation and Development, Paris.

Azoulay, Pierre, Benjamin F. Jones, J. Daniel Kim, and Javier Miranda. 2020. "Age and High-Growth Entrepreneurship." *American Economic Review: Insights* 2 (1): 65–82.

Bajgar, Matej, Giuseppe Berlingieri, Sara Calligaris, Chiara Criscuolo, and Jonathan Timmis. 2020. "Coverage and Representativeness of Orbis Data." OECD Science, Technology and Industry Working Papers No. 6, Organisation for Economic Co-operation and Development, Paris.

Bartelsman, Eric, John C. Haltiwanger, and Stefano Scarpetta. 2004. "Microeconomic Evidence of Creative Destruction in Industrial and Developing Countries." *World Bank Policy Research Working Paper No. 3464*. World Bank, Washington, D.C.

Belenzon, Sharon, Aaron Chatterji, and Brendan Daley. 2017. "Eponymous Entrepreneurs." *American Economic Review* 107 (6): 1638–1655.

Bernard, Andrew, Emmanuel Dhyne, Glenn Magerman, Kalina Manova, and Andreas Moxnes. 2022. "The Origins of Firm Heterogeneity: A Production Network Approach." *Journal of Political Economy* 130 (7): 1765–1804.

Biondi, Filippo, Sergio Inferrera, Matthias Mertens, and Javier Miranda. 2023. "Declining

Job Reallocation in Europe: The Role of Shocks, Market Power, and Technology." *IWH Discussion Paper No. 19-2023*. Halle Institute for Economic Research, Halle (Saale).

Bloom, Nicholas, Rachel Griffith, and John Van Reenen. 2002. "Do R&D Tax Credits Work? Evidence from a Panel of Countries 1979–1997." *Journal of Public Economics* 85 (1): 1–31.

Bolton, Patrick, Neng Wang, and Jinqiang Yang. 2019. "Investment under Uncertainty with Financial Constraints." *Journal of Economic Theory* 184: 104912.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa. 2022. "Discretizing Unobserved Heterogeneity." *Econometrica* 90 (2): 625–643.

Botelho, Tristan L., Daniel C. Fehder, and Yael V. Hochberg. 2026. "Innovation-Driven Entrepreneurship." *Journal of Economic Literature* 64 (1): 89–140.

Brandt, Loren, Gueorgui Kambourov, and Kjetil Storesletten. 2026. "Barriers to Entry and Regional Economic Growth in China." *Review of Economic Studies* 93 (1): 286–326.

Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin. 2011. "Finance and Development: A Tale of Two Sectors." *American Economic Review* 101 (5): 1964–2002.

Buera, Francisco J, and Yongseok Shin. 2017. "Productivity Growth and Capital Flows: The Dynamics of Reforms." *American Economic Journal: Macroeconomics* 9 (3): 147–85.

Cooper, Russell W., and John C. Haltiwanger. 2006. "On the Nature of Capital Adjustment Costs." *Review of Economic Studies* 73 (3): 611–633.

Criscuolo, Chiara, Ralf Martin, Henry G. Overman, and John Van Reenen. 2019. "Some Causal Effects of an Industrial Policy." *American Economic Review* 109 (1): 48–85.

Da Rin, Marco, Marina Di Giacomo, and Alessandro Sembenelli. 2010. "Corporate Taxation and the Size of New Firms: Evidence from Europe." *Journal of the European Economic Association* 8 (2–3): 606–616.

Da Rin, Marco, Marina Di Giacomo, and Alessandro Sembenelli. 2011. "Entrepreneurship, Firm Entry, and the Taxation of Corporate Income: Evidence from Europe." *Journal of Public Economics* 95 (9–10): 1048–1066.

De Loecker, Jan, and Chad Syverson. 2021. "An Industrial Organization Perspective on Productivity." In *Handbook of Industrial Organization*, edited by Kate Ho, Ali Hortacsu, and Alessandro Lizzeri, vol. 4.

Decker, Ryan A., John C. Haltiwanger, Ron S. Jarmin, and Javier Miranda. 2017. "Declining Dynamism, Allocative Efficiency, and the Productivity Slowdown." *American Economic Review* 107 (5): 322–26.

Dent, Robert C, Fatih Karahan, Benjamin Pugsley, and Ayşegül Şahin. 2016. "The Role of Startups in Structural Transformation." *American Economic Review Papers and Proceedings*

106 (5): 219–23.

Derrien, François, Jean-Stéphane Mésonnier, and Guillaume Vuillemey. 2025. "Setup Costs and the Financing of Young Firms." *Management Science*.

Dinlersoz, Emin, Sebnem Kalemli-Özcan, Henry Hyatt, and Veronika Penciakova. 2019. "Leverage over the Firm Life Cycle, Firm Growth, and Aggregate Fluctuations." Working Paper 2019-18, Federal Reserve Bank of Atlanta.

Doms, Mark, and Timothy Dunne. 1998. "Capital Adjustment Patterns in Manufacturing Plants." *Review of Economic Dynamics* 1 (2): 409–429.

Dottori, Davide, and Giacinto Micucci. 2018. "Corporate Liquidity in Italy and Its Increase in the Long Recession." *Economia Politica* 35 (3): 981–1014.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu. 2023. "How Costly are Markups?." *Journal of Political Economy* 131 (7): 1619–1675.

Edwards, Alexander, and Maximilian Todtenhaupt. 2020. "Capital Gains Taxation and Funding for Start-ups." *Journal of Financial Economics* 138 (2): 549–571.

Everitt, Brian S., Sabine Landau, Morven Leese, and Daniel Stahl. 2011. *Cluster Analysis*. 5th ed. Chichester: Wiley.

Ewens, Michael, Kairong Xiao, and Ting Xu. 2024. "Regulatory Costs of Being Public: Evidence from Bunching Estimation." *Journal of Financial Economics* 153: 103775.

Fasano, Francesco, and Marc Deloof. 2021. "Local Financial Development and Cash Holdings in Italian SMEs." *International Small Business Journal* 39 (8): 781–799.

Foster, Lucia, John Haltiwanger, and C.J. Krizan. 2001. "Aggregate Productivity Growth: Lessons from Microeconomic Evidence." In *New Developments in Productivity Analysis*, edited by Edward Dean, Michael Harper and Charles Hulten, 303–372. Chicago: University of Chicago Press.

Garicano, Luis, Claire Lelarge, and John Van Reenen. 2016. "Firm Size Distortions and the Productivity Distribution: Evidence from France." *American Economic Review* 106 (11): 3439–79.

Gopinath, Gita, Şebnem Kalemli-Özcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez. 2017. "Capital Allocation and Productivity in South Europe." *Quarterly Journal of Economics* 132 (4): 1915–1967.

Guner, Nezih, Gustavo Ventura, and Yi Xu. 2008. "Macroeconomic Implications of Size-Dependent Policies." *Review of Economic Dynamics* 11 (4): 721–744.

Gutiérrez, Germán, Callum Jones, and Thomas Philippon. 2021. "Entry Costs and Aggregate Dynamics." *Journal of Monetary Economics* 124: S77–S91.

Guzman, Jorge, and Aishen Li. 2023. "Measuring Founding Strategy." *Management Science* 69 (1): 101–118.

Guzman, Jorge, and Scott Stern. 2015. "Where is Silicon Valley?." *Science* 347 (6222): 606–609.

Haltiwanger, John, Ron S Jarmin, and Javier Miranda. 2013. "Who Creates Jobs? Small versus Large versus Young." *Review of Economics and Statistics* 95 (2): 347–361.

Hopenhayn, Hugo A. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica* 60 (5): 1127–1150.

Hopenhayn, Hugo A., and Richard Rogerson. 1993. "Job Turnover and Policy Evaluation: A General Equilibrium Analysis." *Journal of Political Economy* 101 (5): 915–938.

Hottman, Colin J., Stephen J. Redding, and David E. Weinstein. 2016. "Quantifying the Sources of Firm Heterogeneity." *Quarterly Journal of Economics* 131 (3): 1291–1364.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics* 124 (4): 1403–1448.

Hsieh, Chang-Tai, and Peter J. Klenow. 2014. "The Life Cycle of Plants in India and Mexico." *Quarterly Journal of Economics* 129 (3): 1035–1084.

Hurst, Erik, and Benjamin W. Pugsley. 2011. "What Do Small Businesses Do?." *Brookings Papers on Economic Activity* 43 (2): 73–142.

Juhász, Réka, Nathan J Lane, and Dani Rodrik. 2024. "The New Economics of Industrial Policy." *Annual Review of Economics* 16: 213–242.

Kalemli-Özcan, Şebnem, Bent E Sørensen, Carolina Villegas-Sanchez, Vadym Volosovych, and Sevcan Yeşiltaş. 2024. "How to Construct Nationally Representative Firm-Level Data from the Orbis Global Database: New Facts on SMEs and Aggregate Implications for Industry Concentration." *American Economic Journal: Macroeconomics* 16 (2): 353–374.

Kalra, Aarti. 2019. "Understanding the Corporate Tax Structure for Domestic Startup Companies." Startup India, Government of India.

Karahan, Fatih, Benjamin Pugsley, and Ayşegül Şahin. 2024. "Demographic Origins of the Startup Deficit." *American Economic Review* 114 (7): 1986–2023.

Kerr, William R., and Ramana Nanda. 2010. "Banking Deregulations, Financing Constraints, and Firm Entry Size." *Journal of the European Economic Association* 8 (2-3): 582–593.

Kochen, Federico. 2025. "Finance over the Life Cycle of Firms." *CEMFI Working Paper No. 2518*. CEMFI, Madrid.

Larrain, Mauricio, and Sebastian Stumpner. 2017. "Capital Account Liberalization and

Aggregate Productivity: The Role of Firm Capital Allocation." *Journal of Finance* 72 (4): 1825–1858.

Lemmon, Michael L., Michael R. Roberts, and Jaime F. Zender. 2008. "Back to the Beginning: Persistence and the Cross-section of Corporate Capital Structure." *Journal of Finance* 63 (4): 1575–1608.

Levine, Ross, and Yona Rubinstein. 2017. "Smart and Illicit: Who Becomes an Entrepreneur and Do They Earn More?." *Quarterly Journal of Economics* 132 (2): 963–1018.

Lopez-Garcia, Paloma, and Filippo di Mauro. 2015. "Assessing European Competitiveness: The New CompNet Microbased Database." Working Paper 1764, European Central Bank.

Ma, Song, Justin Murfin, and Ryan Pratt. 2022. "Young Firms, Old Capital." *Journal of Financial Economics* 146 (1): 331–356.

Mian, Atif, and Amir Sufi. 2014. "What Explains the 2007–2009 Drop in Employment?." *Econometrica* 82 (6): 2197–2223.

Midrigan, Virgiliu, and Daniel Yi Xu. 2014. "Finance and Misallocation: Evidence from Plant-Level Data." *American Economic Review* 104 (2): 422–58.

Moll, Benjamin. 2014. "Productivity Losses from Financial Frictions: Can Self-financing undo Capital Misallocation?." *American Economic Review* 104 (10): 3186–3221.

Moreira, Sara. 2017. "Firm Dynamics, Persistent Effects of Entry Conditions, and Business Cycles." Working Paper CES-WP-17-29, U.S. Census Bureau, Center for Economic Studies.

Oberfield, Ezra, and Devesh Raval. 2021. "Micro Data and Macro Technology." *Econometrica* 89 (2): 703–732.

OECD. 2026a. "Corporate Income Tax Rates Database: Statutory and Targeted Small Business Corporate Income Tax Rates." `https://www.oecd.org/en/data/datasets/corporate-income-tax-rates-database.html`.

OECD. 2026b. *OECD Compendium of Productivity Indicators 2026*. Paris: OECD Publishing.

PwC. 2026. "Spain: Corporate – Taxes on Corporate Income." Worldwide Tax Summaries, `https://taxsummaries.pwc.com/spain/corporate/taxes-on-corporate-income`.

Restuccia, Diego, and Richard Rogerson. 2013. "Misallocation and Productivity." *Review of Economic Dynamics* 16 (1): 1–10.

Robb, Alicia M., and David T. Robinson. 2014. "The Capital Structure Decisions of New Firms." *Review of Financial Studies* 27 (1): 153–179.

Schoar, Antoinette. 2010. "The Divide between Subsistence and Transformational Entrepreneurship." *Innovation Policy and the Economy* 10 (1): 57–81.

Sedláček, Petr, and Vincent Sterk. 2017. "The Growth Potential of Startups over the Business Cycle." *American Economic Review* 107 (10): 3182–3210.

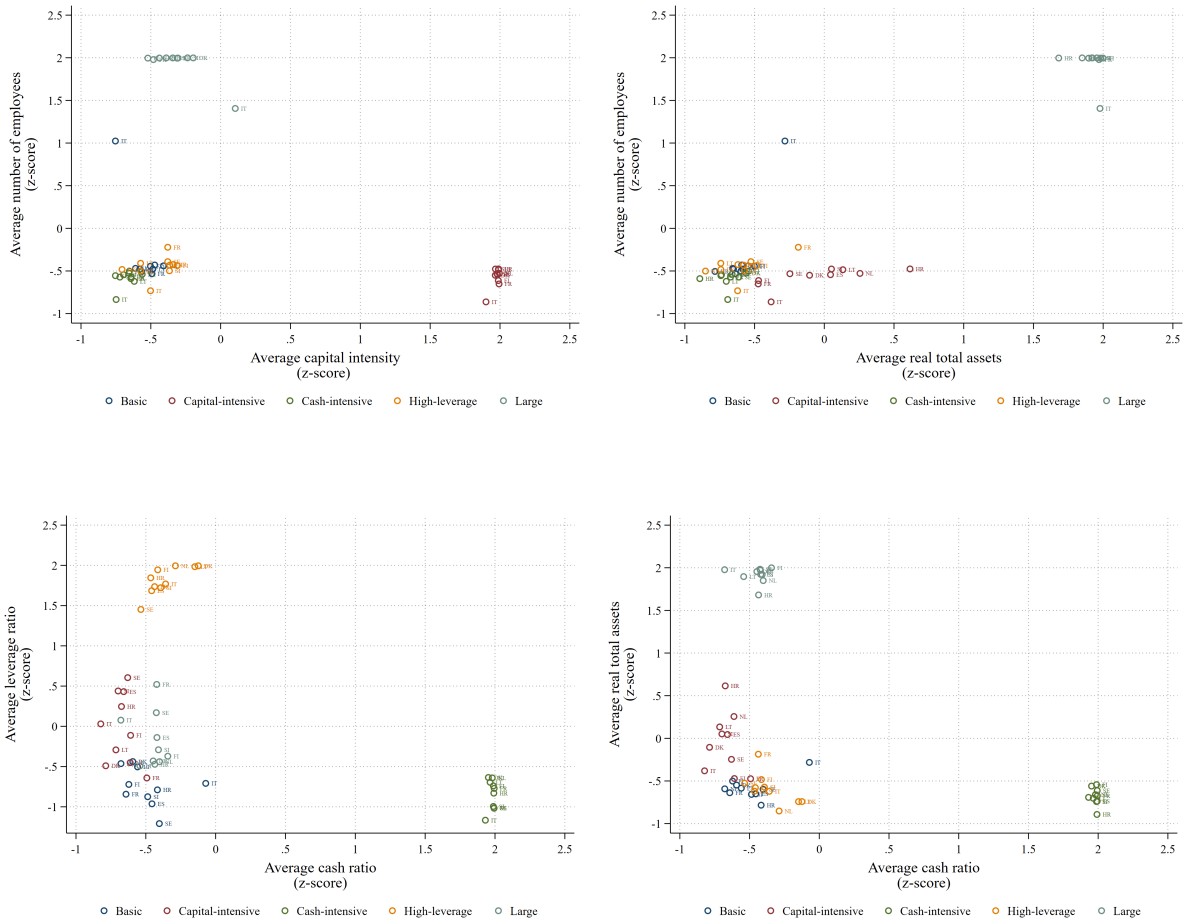
Sterk, Vincent ⓡ Sedláček, Petr ⓡ Benjamin W. Pugsley. 2021. "The Nature of Firm Growth." *American Economic Review* 111 (2): 547–579.

Syverson, Chad. 2011. "What Determines Productivity?." *Journal of Economic Literature* 49 (2): 326–365.

Townsend, Robert M. 1979. "Optimal Contracts and Competitive Markets with Costly State Verification." *Journal of Economic Theory* 21 (2): 265–293.

Tables and Figures

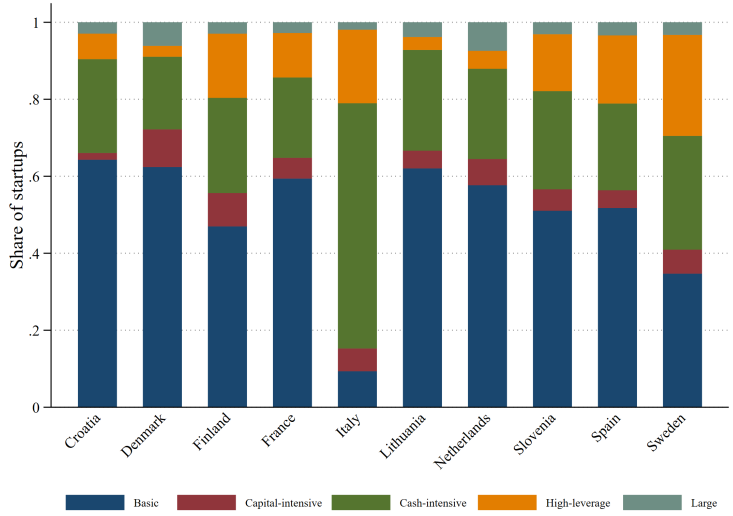
Figure 1. Meta-clustering of startup types



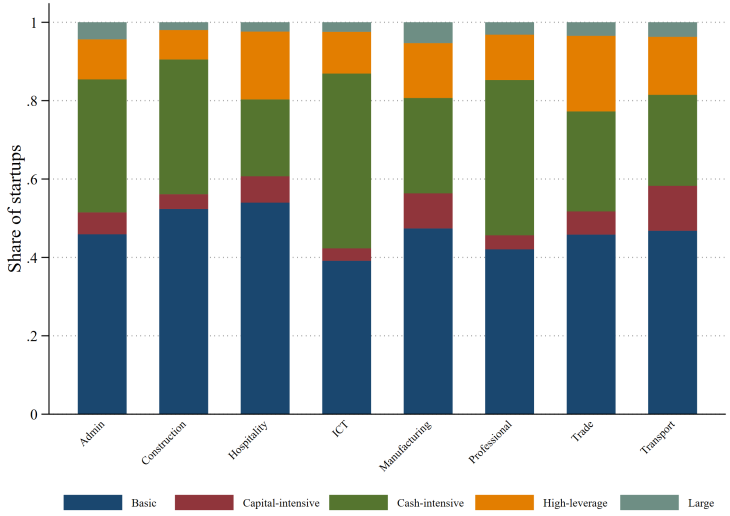
Notes: The four panels in this figure summarize the meta-clustering procedure. Different meta-clusters are denoted by different colors. The meta-clustering groups comparable clusters from different countries, taking the first-stage cluster centers derived from each country's clustering procedure as the observations. More specifically, in the meta-clustering procedure, the units of observation are z-scores of the first-stage cluster centers, averaged across years and industries.

Figure 2. Distribution of startup types by country, industry, and cohort

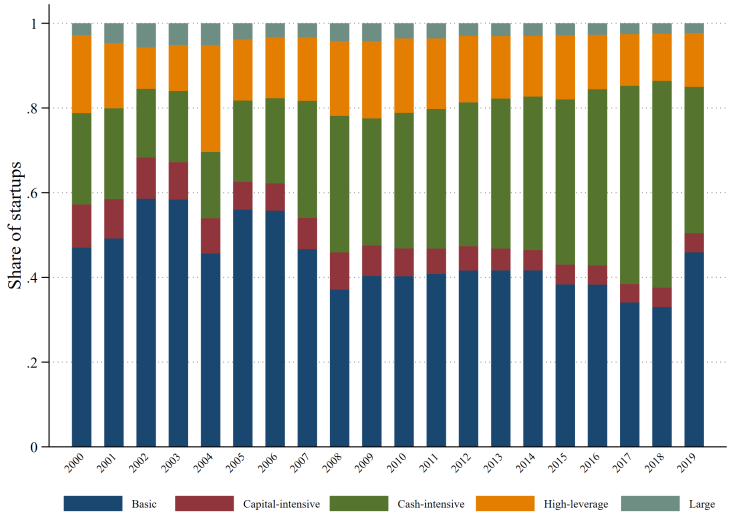
Panel A: Distribution of startup types by country



Panel B: Distribution of startup types by one-digit industry

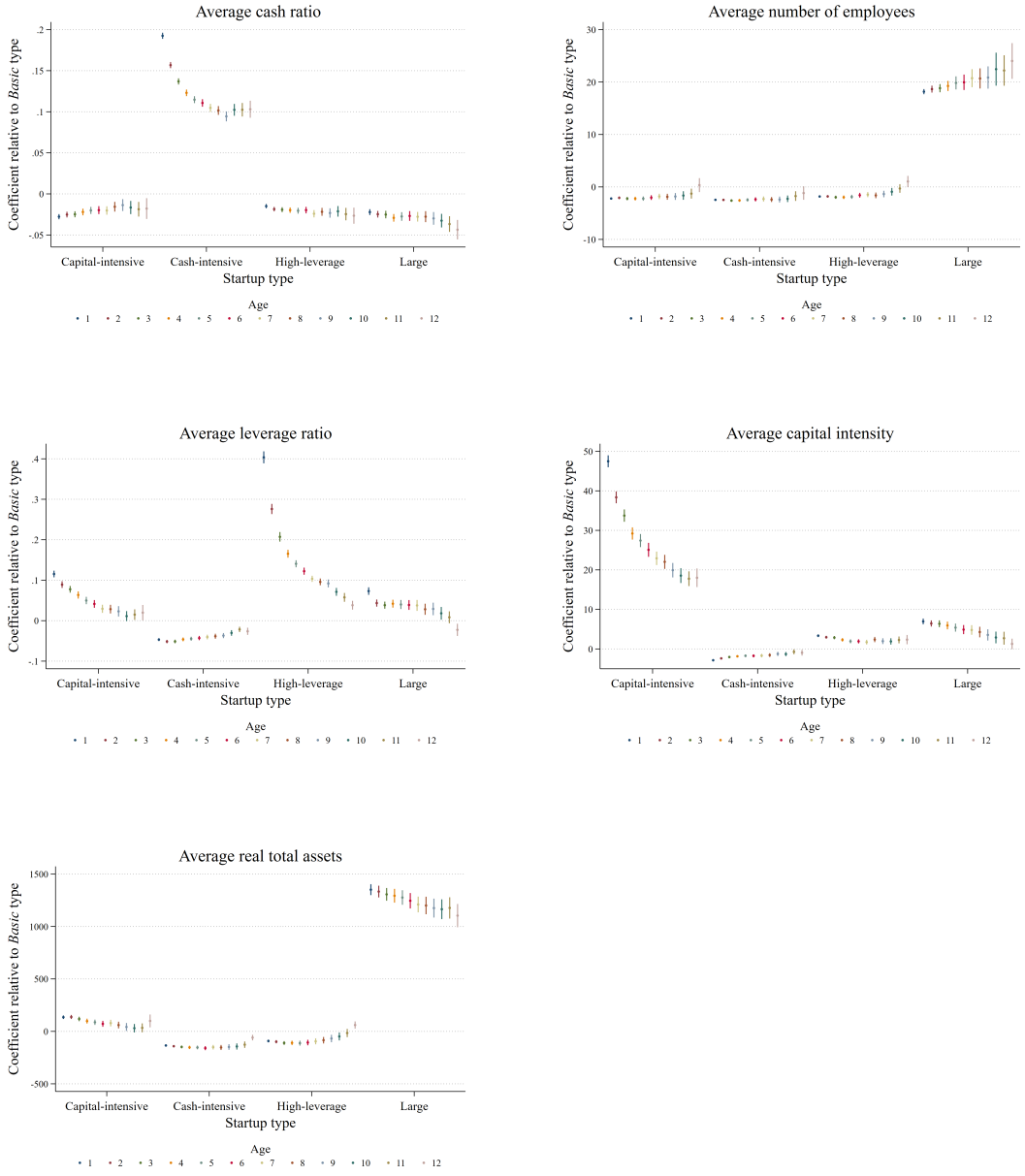


Panel C: Distribution of startup types by cohort



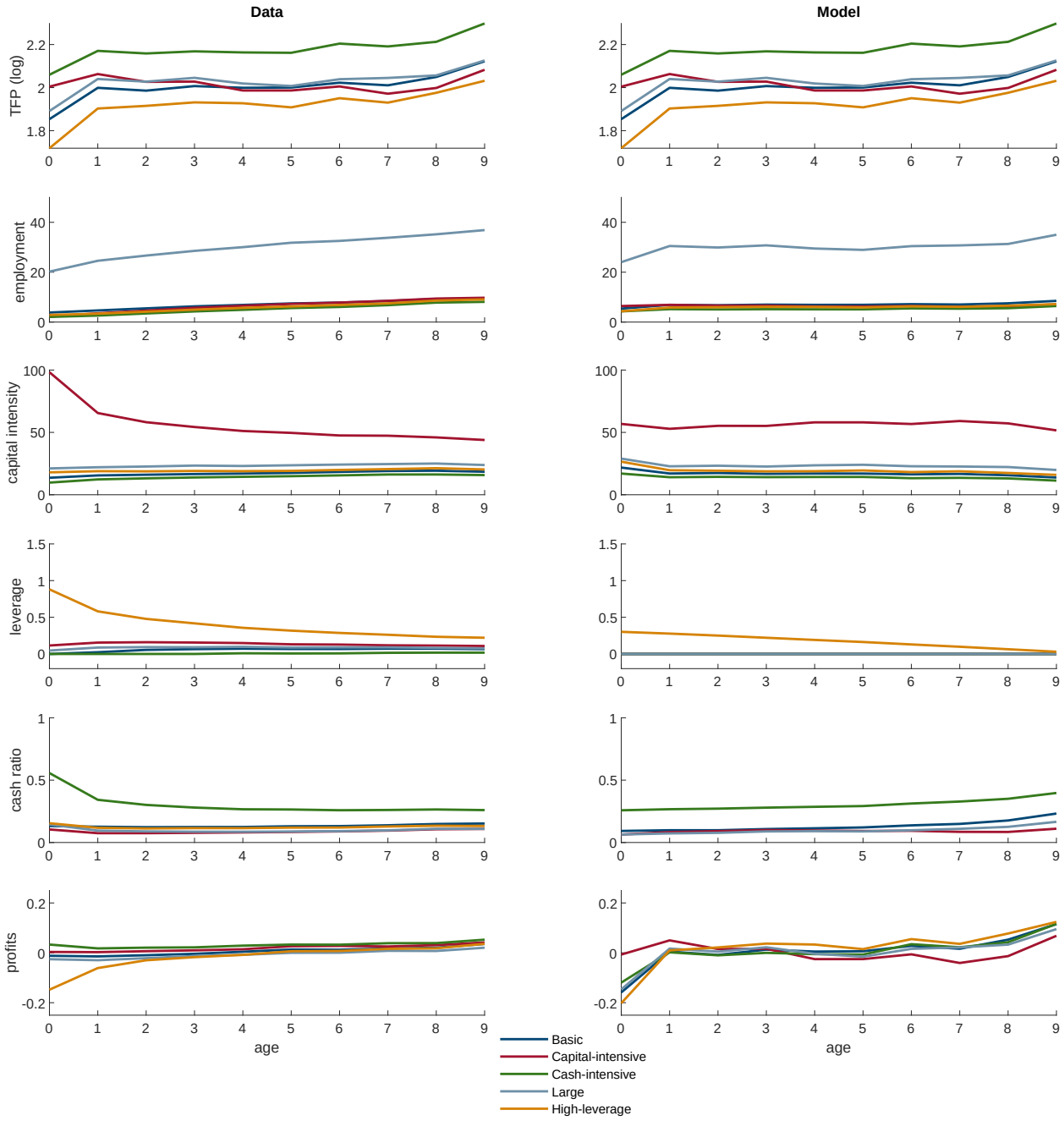
Notes: This figure illustrates the distribution of the startup population across the five startup types for individual countries (Panel A), one-digit NACE Rev.2 industries (Panel B), and cohorts (Panel C). The startup population comprises firms in their year of entry (age 0) in all cohorts available for each country.

Figure 3. The life cycle of startup types



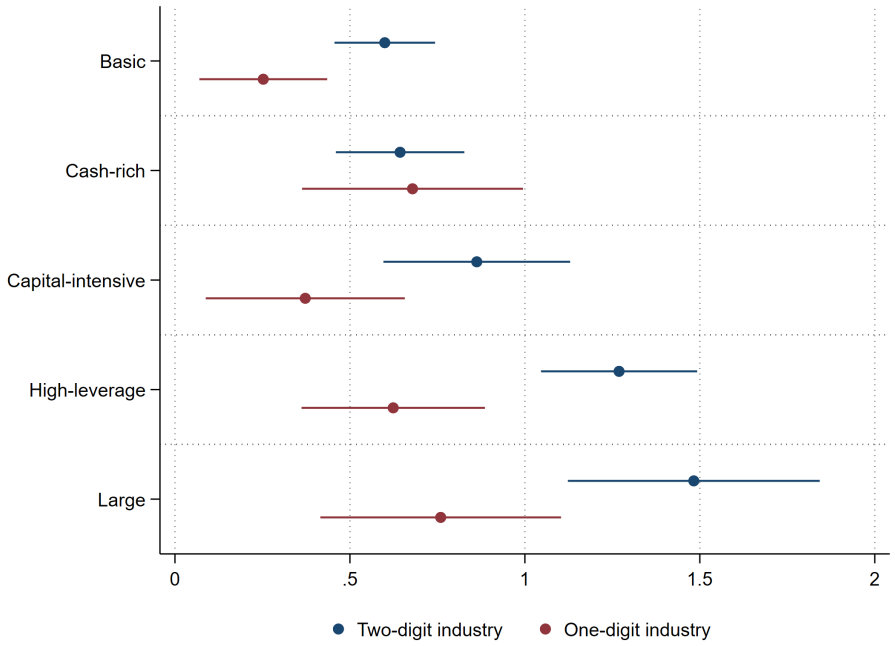
Notes: The panels in this figure summarize how the startup types develop during the first 12 years of their life in terms of the five clustering variables. Each panel corresponds to one clustering variable and plots the coefficients from 12 separate regressions where the dependent variable is this clustering variable. Each regression is then run for each age (age is 1, 2, ...,12 years). For example, the first panel summarizes regressions in which the *Average cash to total assets* ratio is regressed on dummy variables for the startup types (the *Basic* type is omitted) as well as *Country* × *Cohort* and *Industry* fixed effects. The sample is the full panel dataset at the two-digit industry level. Standard errors are clustered at the *Industry* × *Country* × *Cohort* level, and whiskers indicate 95 percent confidence intervals. Capital intensity is measured in thousands of euros (PPP-adjusted) per employee; the cash and leverage ratios are expressed relative to total assets.

Figure 4. Startup profiles: Data versus model



Notes: See the main text and appendix for a description of the model and calibration procedure. The charts on the left show unweighted averages across countries and cohorts. Starting from data aggregated to the country-startup type-cohort-age level, we regress outcomes on *Type* × *Age FE* and *Country* × *Age FE*. For each outcome, we plot the total effect for each type at each age.

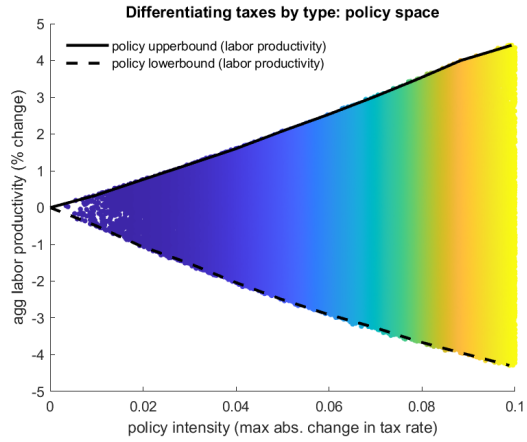
Figure 5. (Inverse) entry elasticities by startup type



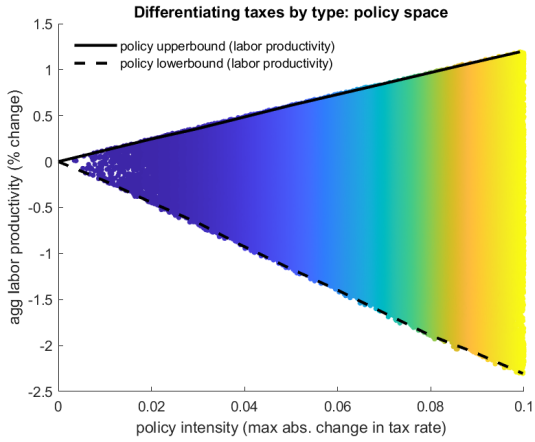
Notes: This figure shows coefficients from two OLS regressions where the dependent variable is the log net present value of firm profits. The estimated coefficients represent the inverse entry elasticity, β_1, for each startup type. *Country* $\times$ *Startup type* $\times$ *Industry* and *Country* $\times$ *Industry* $\times$ *Cohort* fixed effects are included throughout. Blue and red markers depict regressions at the two-digit and one-digit NACE Rev.2 industry level, respectively. The dataset only includes cohorts observed for more than seven years (those founded in or before 2011). For age 12 and onward, profits and the year-on-year exit rate are assumed fixed. Standard errors are clustered at the *Industry* $\times$ *Country* $\times$ *Cohort* level, and whiskers indicate 95 percent confidence intervals.

Figure 6. Policy experiment: Differentiation of corporate income tax

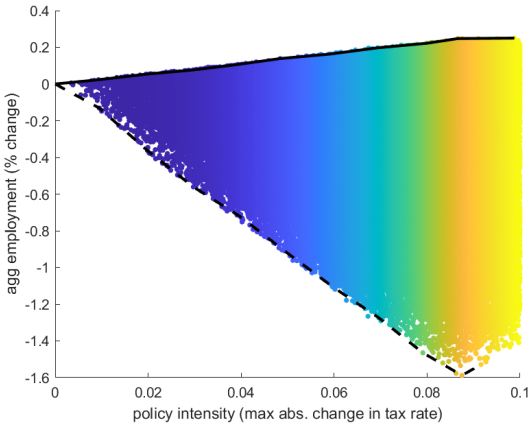
Panel A: Labor productivity (one-digit)



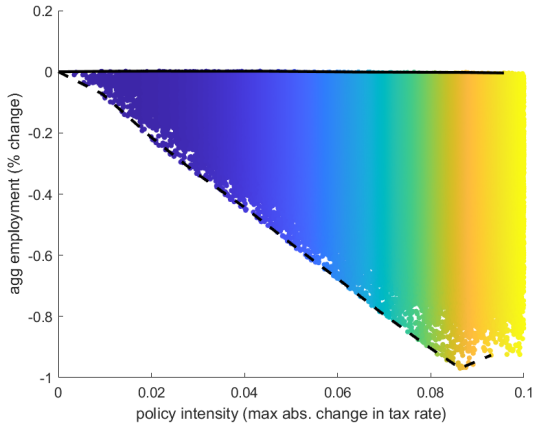
Panel B: Labor productivity (two-digit)



Panel C: Employment (one-digit)

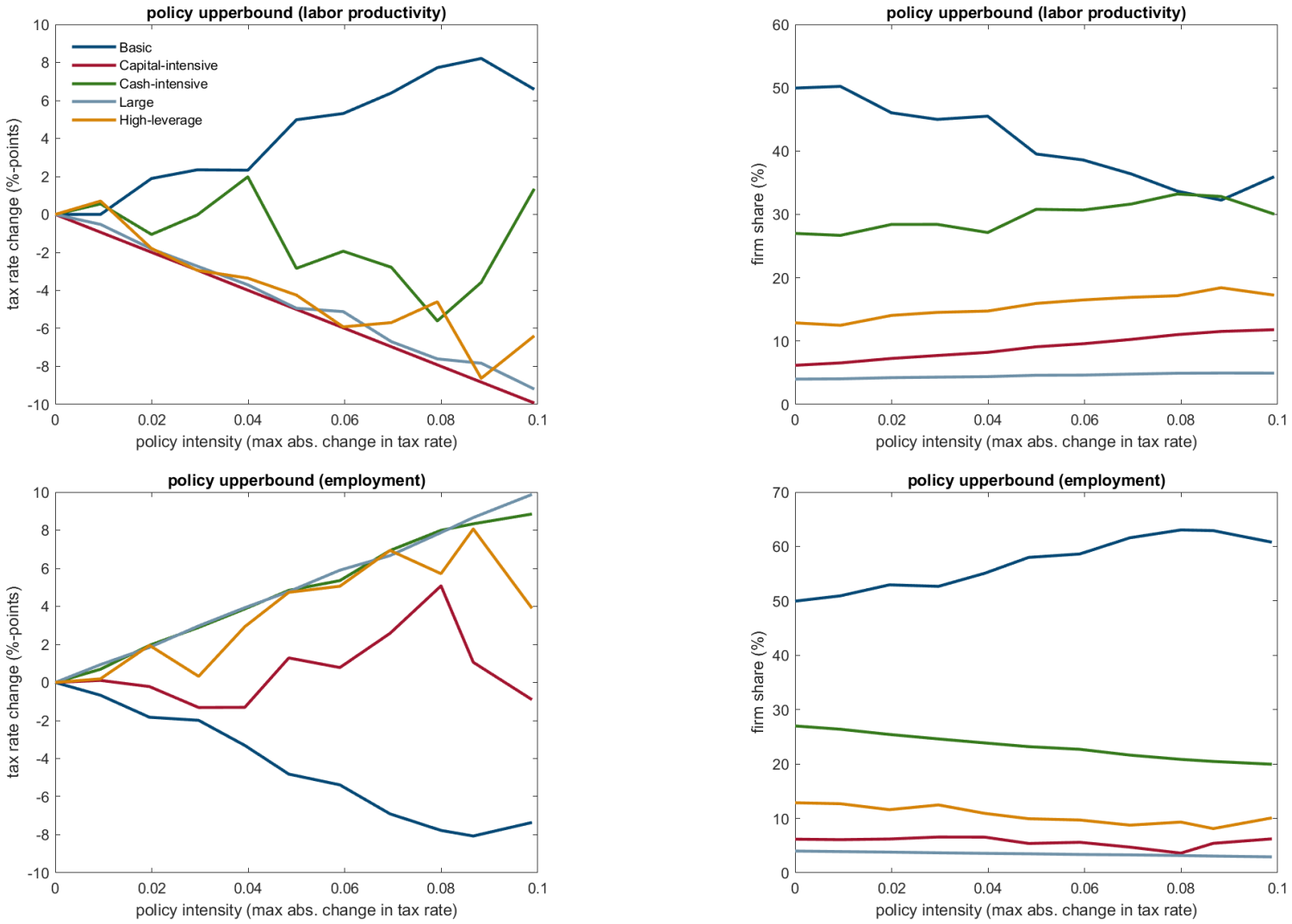


Panel D: Employment (two-digit)



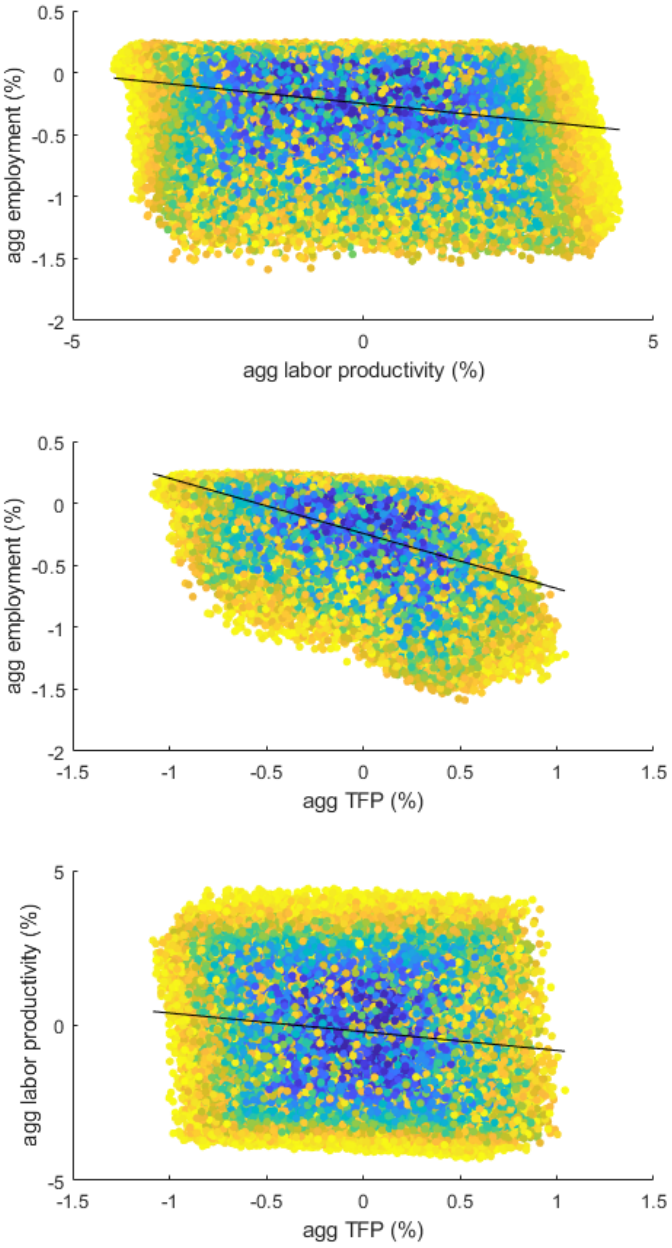
Notes: This figure summarizes the differentiated corporate income tax experiment. The top row (Panels A and B) shows the policy space for aggregate labor productivity; the bottom row (Panels C and D) shows it for aggregate employment. The left column (Panels A and C) uses entry elasticities based on one-digit industry data; the right column (Panels B and D) uses two-digit industry data. In each panel the horizontal axis measures the intensity of the potential policy change, given by the maximum absolute change in the tax rate across types, and warmer colors indicate stronger corporate tax rate differentiation. The solid line plots the "policy upper bound", the largest possible change in the outcome given a certain policy intensity, and the dashed line plots the corresponding "policy lower bound".

Figure 7. Policy experiment: Upper bounds of macro impacts due to tax differentiation



Notes: This figure shows the specific tax policies and resulting startup shares associated with the two policy upper bounds from Figure 6, Panels A and C (one-digit sector classification). The two panels on the left (right) plot the corporate tax rate policies (startup shares) associated with the two policy upper bounds. The upper left (right) panel shows the tax rates (startup shares) associated with the labor productivity upper bound. The lower left (right) panel shows the tax rates (startup shares) associated with the employment upper bound.

Figure 8. Policy experiment: Macroeconomic trade-offs



Notes: These scatter plots each depict pairs of changes in aggregate or average macroeconomic outcomes resulting from each potential policy (up to a policy intensity of 0.1), using the one-digit sector classification. Linear regression lines are shown in black. Warmer colors indicate higher policy intensity, as in Figure 6.

Table 1. Characteristics of startup types at time of entry

	(1)	(2)	(3)	(4)	(5)	(6)
Startup type	No. of employees	Capital intensity	Real total assets	Cash ratio	Leverage ratio	Share of startups (%)
Basic	4	8.56	166.59	0.12	0.23	49
Capital-intensive	2	93.18	405.12	0.09	0.41	7
Cash-intensive	2	4.67	92.46	0.54	0.18	26
High-leverage	3	12.90	122.97	0.14	1.18	14
Large	20	16.05	1488.30	0.13	0.34	4

Notes: This table presents the cross-country means of the cluster variables for each of the five startup types in the year of establishment. Means are unweighted and based on all available cohorts and countries. Capital intensity is measured in thousands of euros of fixed assets per employee, and real total assets in thousands of euros (PPP-adjusted). Column (6) reports each type's share of all startups in the sample.

Table 2. Startup types and firm outcomes

Panel A: All firms, aged 0-9

	(1)	(2)	(3)	(4)	(5)
	Aggregate labor productivity	Aggregate TFP	Average exit probability	Average wage per employee	Average profit margin
Capital-intensive	0.308***	0.047***	-0.061***	2.359***	0.012***
	(0.004)	(0.004)	(0.003)	(0.091)	(0.001)
Cash-intensive	0.030***	0.050***	-0.009***	1.104***	0.022***
	(0.003)	(0.002)	(0.002)	(0.056)	(0.001)
High-leverage	-0.044***	-0.038***	-0.000	-1.494***	-0.030***
	(0.003)	(0.003)	(0.002)	(0.062)	(0.001)
Large	0.165***	0.045***	-0.125***	3.391***	-0.018***
	(0.004)	(0.003)	(0.003)	(0.092)	(0.001)
Constant	3.362***	2.216***	0.717***	27.904***	0.042***
	(0.002)	(0.001)	(0.001)	(0.032)	(0.000)
R-squared	0.904	0.978	0.639	0.910	0.604
N	28,991	20,786	31,240	30,259	29,992

Panel B: All firms, aged 5-9

	(1)	(2)	(3)	(4)	(5)
	Aggregate labor productivity	Aggregate TFP	Average exit probability	Average wage per employee	Average profit margin
Capital-intensive	0.190***	0.011**	-0.062***	0.832***	0.013***
	(0.007)	(0.006)	(0.003)	(0.114)	(0.001)
Cash-intensive	0.018***	0.031***	-0.024***	1.174***	0.013***
	(0.005)	(0.004)	(0.003)	(0.088)	(0.001)
High-leverage	-0.025***	-0.022***	0.001	-1.328***	-0.008***
	(0.005)	(0.005)	(0.002)	(0.097)	(0.001)
Large	0.151***	0.048***	-0.140***	2.669***	-0.013***
	(0.007)	(0.006)	(0.003)	(0.144)	(0.001)
Constant	3.473***	2.251***	0.817***	30.289***	0.049***
	(0.003)	(0.003)	(0.001)	(0.048)	(0.001)
R-squared	0.904	0.981	0.618	0.929	0.648
N	10,160	6,993	11,167	10,664	10,583
Country × Cohort FE	✓	✓	✓	✓	✓
Industry × Cohort FE	✓	✓	✓	✓	✓
Country × Industry FE	✓	✓	✓	✓	✓
Age × Country FE	✓	✓	✓	✓	✓
Age × Industry FE	✓	✓	✓	✓	✓
Age × Cohort FE	✓	✓	✓	✓	✓

Notes: This table shows OLS regressions where the dependent variable is indicated in the column heading. Observations are at the country × sector × startup type × cohort × age level. The dependent variables in the first two columns are aggregate outcomes (i.e. employment-weighted averages) by startup type. The dependent variables in the last three columns are simple averages by startup type. Industries are defined at the one-digit NACE Rev.2 level. Regressions are based on the full panel of firms aged 0-9 in Panel A and firms aged 5-9 in Panel B. Robust standard errors are in parentheses. Wages are denominated in thousands of euros.

Table 3. Type-specific parameters in the post-entry model

	α	γ	$\alpha + \gamma$	m_{-1}/w	c/w
	Capital elasticity	Labor elasticity	Scalability	Initial equity	Fixed cost
Basic	0.29	0.41	0.70	35.5	6.8
Capital-intensive	0.44	0.18	0.62	100.8	19.4
Cash-intensive	0.24	0.41	0.65	26.6	5.5
High-leverage	0.31	0.38	0.70	21.5	6.2
Large	0.37	0.38	0.75	201.1	31.2

Notes: See the main text and appendix for a description of the model and calibration procedure. The initial equity and fixed costs are scaled by the annual wage per worker for ease of interpretation.

Table 4. Macroeconomic effects of capital, labor, and interest subsidies

	Baseline	Capital subsidy		Labor subsidy			Interest subsidy		
Panel A. Policy rates (%)									
τ_k	0.00	−0.29		0.00			0.00		
τ_l	0.00	0.00		−1.68			0.00		
τ_b	0.00	0.00		0.00			−0.29		
τ_π	0.00	5.00		5.00			0.02		
Panel B. Startup shares (percentage-point change vs. baseline)									
Basic		−0.57		0.54			−0.33		
Capital-intensive		0.62		−1.51			−0.02		
Cash-intensive		−0.10		1.11			−0.06		
Large		0.09		0.03			0.02		
High-leverage		−0.06		−0.17			0.42		
Panel C. Aggregate effects (% change vs. baseline)									
	Total	Comp.	# firms	Total	Comp.	# firms	Total	Comp.	# firms
Output (Y)	0.67	0.82	−1.29	−3.10	−1.25	−3.13	0.03	−0.06	0.07
Capital (K)	2.65	1.33	−1.29	−0.29	−2.35	−3.13	0.04	−0.04	0.07
Employment (L)	0.05	0.25	−1.29	−0.01	0.10	−3.13	0.01	−0.06	0.07
Labor prod. (Y/L)	0.63	0.57	0.00	−3.09	−1.35	0.00	0.02	0.00	0.00
# firms	−1.29	0.00	−1.29	−3.13	0.00	−3.13	0.07	0.00	0.07

Notes: This table reports the effects of three alternative fiscal interventions financed through a uniform adjustment in the corporate income tax rate τ_π, with wages adjusting to clear the labor market. The capital subsidy τ_k is a flow subsidy per unit of installed capital; the labor subsidy τ_l reduces the effective wage paid by the firm; and the interest subsidy τ_b reduces the interest rate on long-term debt and is restricted to firms below age ten. Each intervention is scaled so that the financing requirement corresponds to a five-percentage-point increase in τ_π, except for the interest subsidy, whose total cost is small because of the age restriction. Panel A reports the policy rates implied by each intervention, with all rates expressed in percent. Panel B reports the resulting shifts in the share of each startup type, in percentage points relative to baseline. Panel C reports aggregate effects in percent relative to baseline: "Total" is the overall effect; "Comp." is the contribution from changes in startup composition holding the number of firms and each firms' choices fixed; and "# firms" is the contribution from changes in the total number of firms holding composition and each firm's choices fixed. The remainder (Total minus Comp. minus # firms) reflects the within-type response of firms to the instrument and to the equilibrium wage, together with interaction terms. The baseline is the calibrated economy with zero subsidies and a zero corporate income tax rate.

Internet Appendix for

Startup Types and Macroeconomic Performance in Europe

Ralph De Haas
EBRD, KU Leuven & CEPR

Vincent Sterk
UCL & CEPR

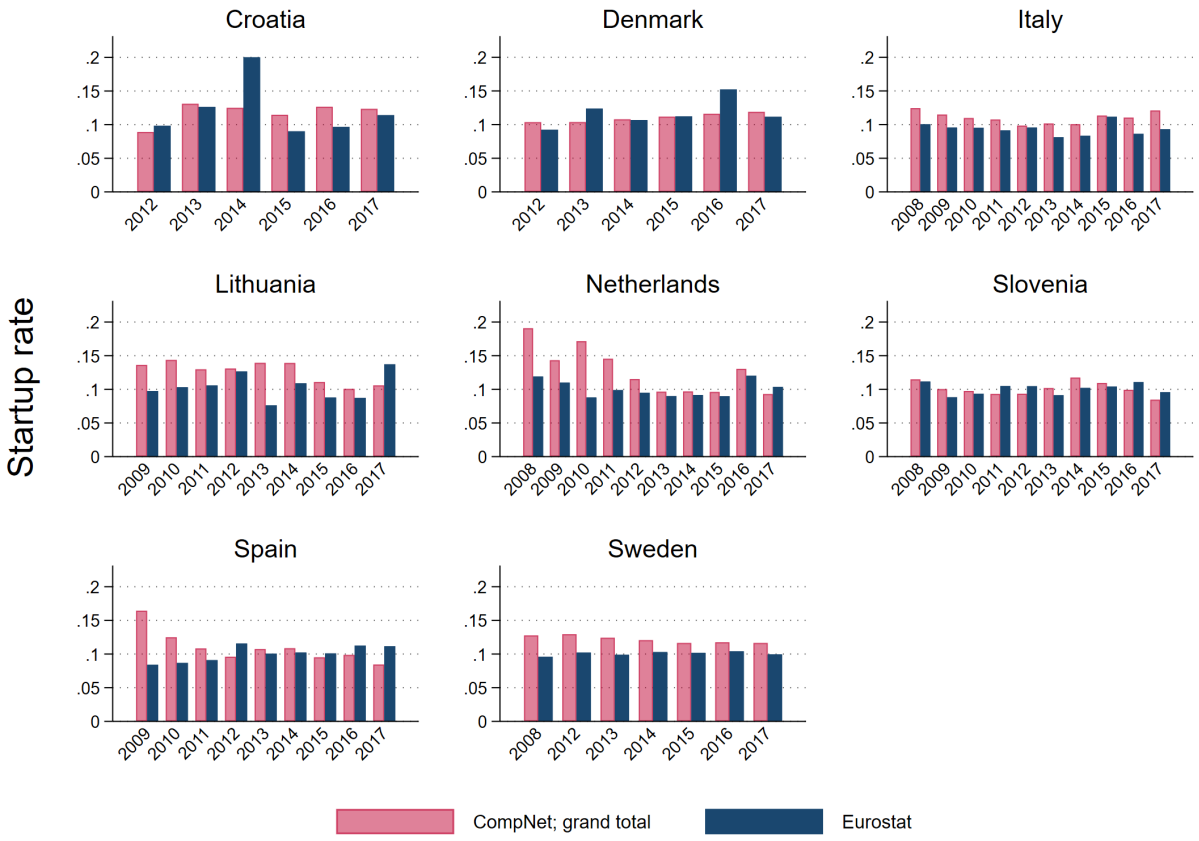
Neeltje Van Horen
University of Amsterdam & CEPR

Internet Appendix A: Additional Figures and Tables

Comparison to Eurostat. In Figures A1–A4, we validate our data by comparing our startup population to Eurostat's Business Demography Statistics (BDS) on startups (while excluding sole proprietorships for consistency). Eurostat is the statistical office of the European Union (EU) and collects the BDS from the national statistical institutes of EU member states. Figure A1 shows that startup rates (the number of startups in a year as a fraction of the total firm population) are very comparable between our CompNet-based dataset and the aggregate data published by Eurostat. The same holds, by and large, when we compare average employment growth during the first five years after startup (Figure A2). In most countries, trend growth in both datasets is very similar. In a few cases—such as the Netherlands and Sweden—there are gaps in the average *level* of reported employees. In those cases, a comparison with other countries suggests the Eurostat data are anomalous, rather than the CompNet data. Next, Figure A3 compares exit rates (firm death) over time. The five-year cumulative exit rate is comparable at 45 percent (56 percent) in the CompNet (Eurostat) dataset. Moreover, the trends as firms age are very similar in both cases, although in Sweden, CompNet tends to overreport firm exit.

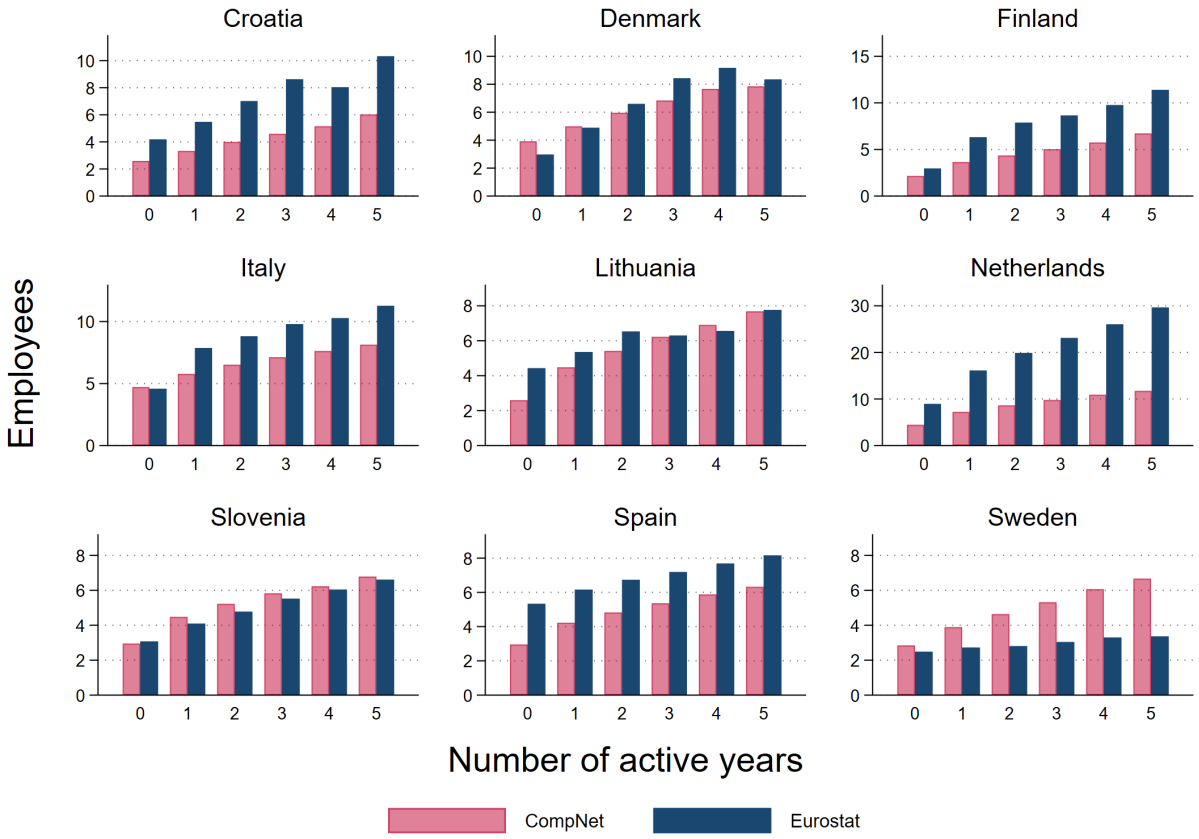
Finally, Figure A4 compares the sectoral composition of startups. We focus on the number of startups in each one-digit industry as a share of total startups across all industries, using pooled data over 2008–2018. While there is cross-country heterogeneity in the sectoral composition of startups, this national composition is very similar when using CompNet or Eurostat data. Overall, we conclude that the firm population that underpins the statistical moments we derive from CompNet is representative of the firm population in our sample countries.

Figure A1. Startup rates by cohort and country: CompNet versus Eurostat



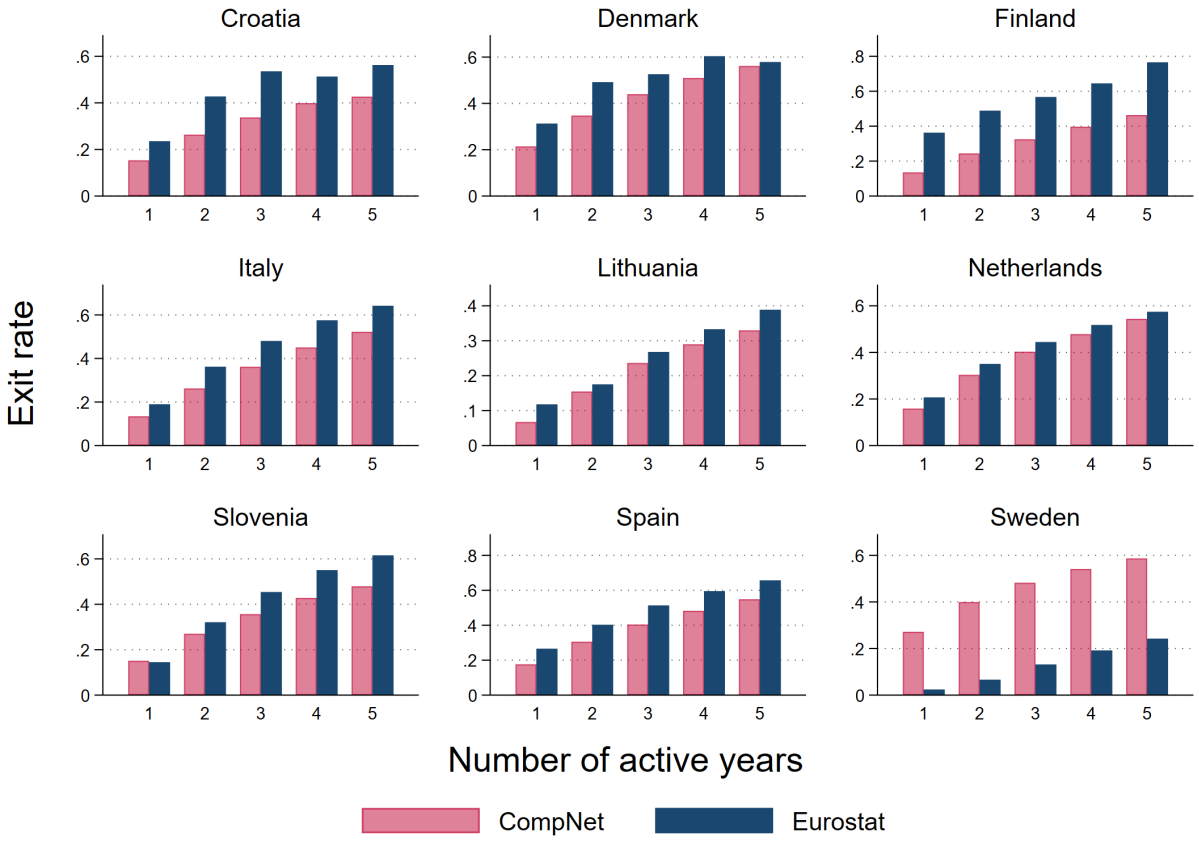
Notes: This figure compares the total number of startups in the CompNet database (pink bars) with establishment of firms in Eurostat (blue bars). Cohorts reported are subject to data availability. Finland is omitted in this chart because we cannot access the total number of startups in the CompNet database for that country. France does not appear as Eurostat does not report firm entry prior to 2008.

Figure A2. Employment by firm age: CompNet versus Eurostat



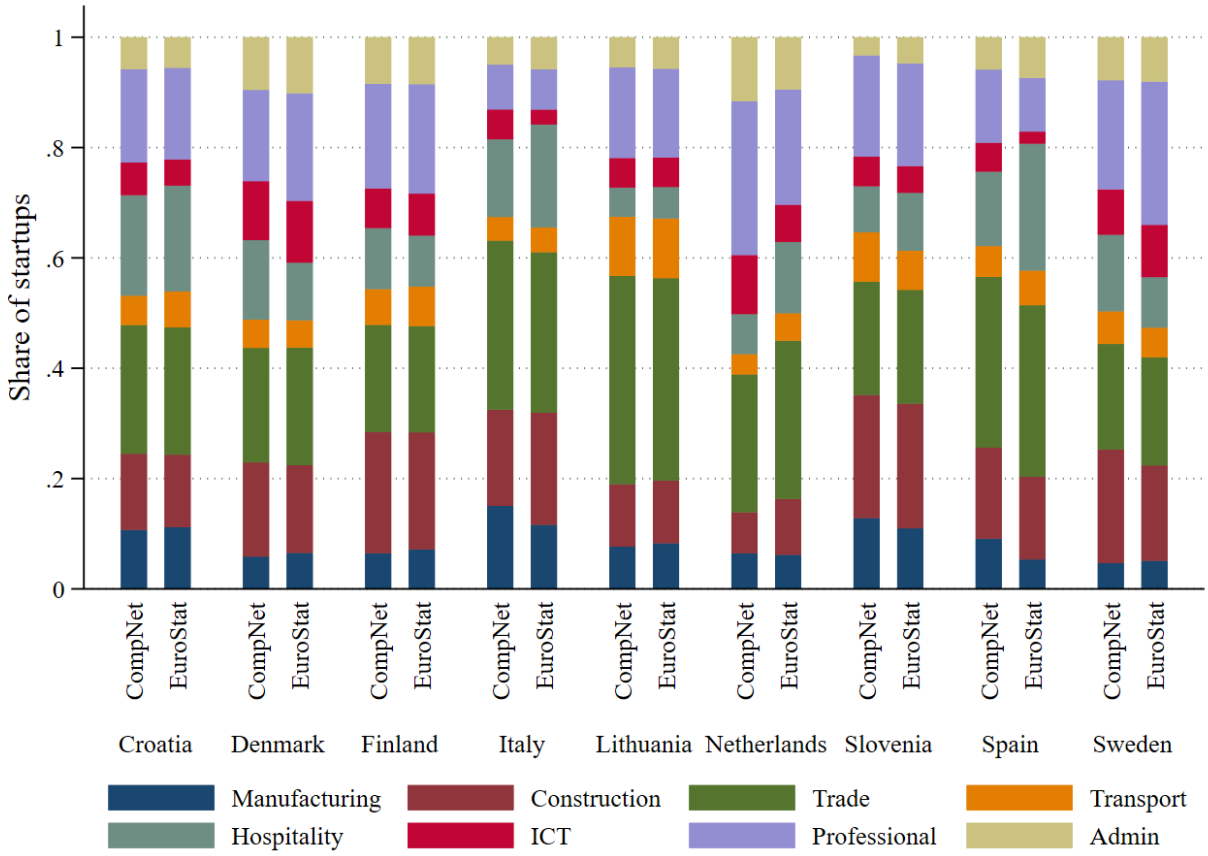
Notes: This figure compares growth in number of persons employed by startups as reported by CompNet (pink bars) and Eurostat (blue bars). CompNet and Eurostat data are matched based on the startup cohort and age group. For comparison purposes, we adjust the Eurostat data such that sole proprietorship firms are removed and we adjust for the average number of persons employed by sole proprietorship firms. The x-axis depicts *Startup age*, which is the number of years a startup has been active. *Employees* on the y-axis is averaged over cohorts. France does not appear as Eurostat does not report firm entry prior to 2008.

Figure A3. Cumulative exit rates by firm age: CompNet versus Eurostat



Notes: This figure compares the cumulative exit rate of startups in CompNet (pink bars) and Eurostat (blue bars). CompNet and Eurostat data are matched based on the startup cohort and age group. The x-axis depicts *Startup age*, which is the number of years a startup has been active. *Exit rate* is the average cumulative exit rate over all cohorts for each startup age group. France does not appear as Eurostat does not report firm entry prior to 2008.

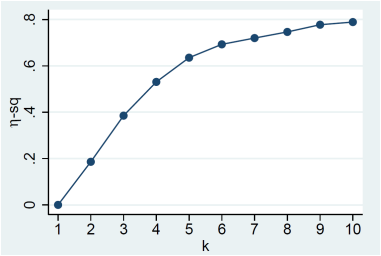
Figure A4. Startup shares by industry: CompNet versus Eurostat



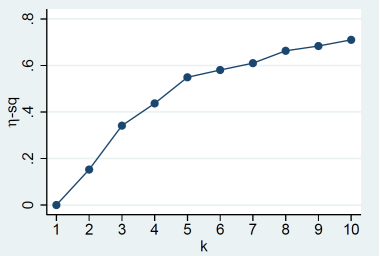
Notes: This figure compares the number of startups in each industry as a share of total startups across the industries. The data is pooled over the years 2008–2018; country × year × industry cells only enter the calculations if the cell is non-missing in both CompNet and Eurostat.

Figure A5. Scree plots used to choose the number of clusters

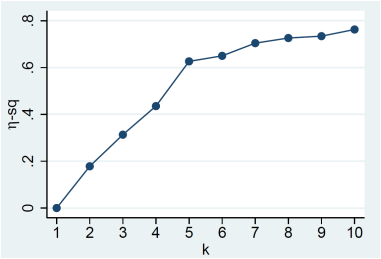
Croatia



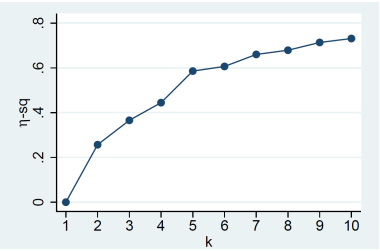
Denmark



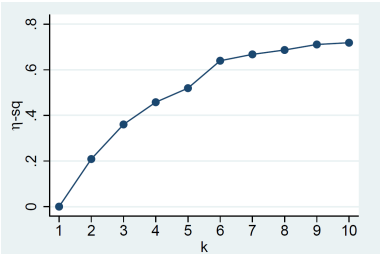
Finland



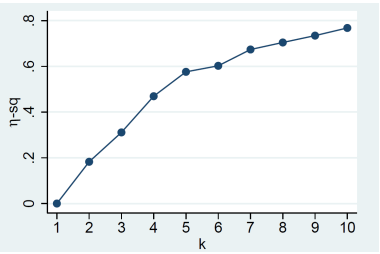
France



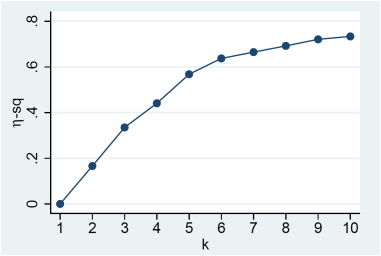
Italy



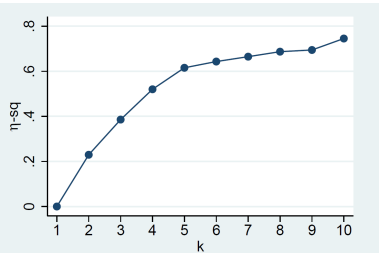
Lithuania



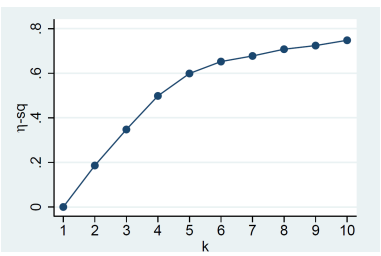
Netherlands



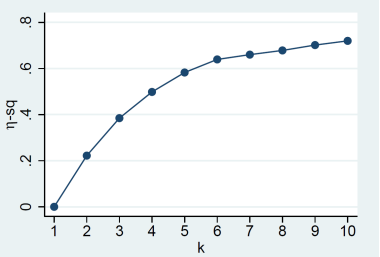
Slovenia



Spain

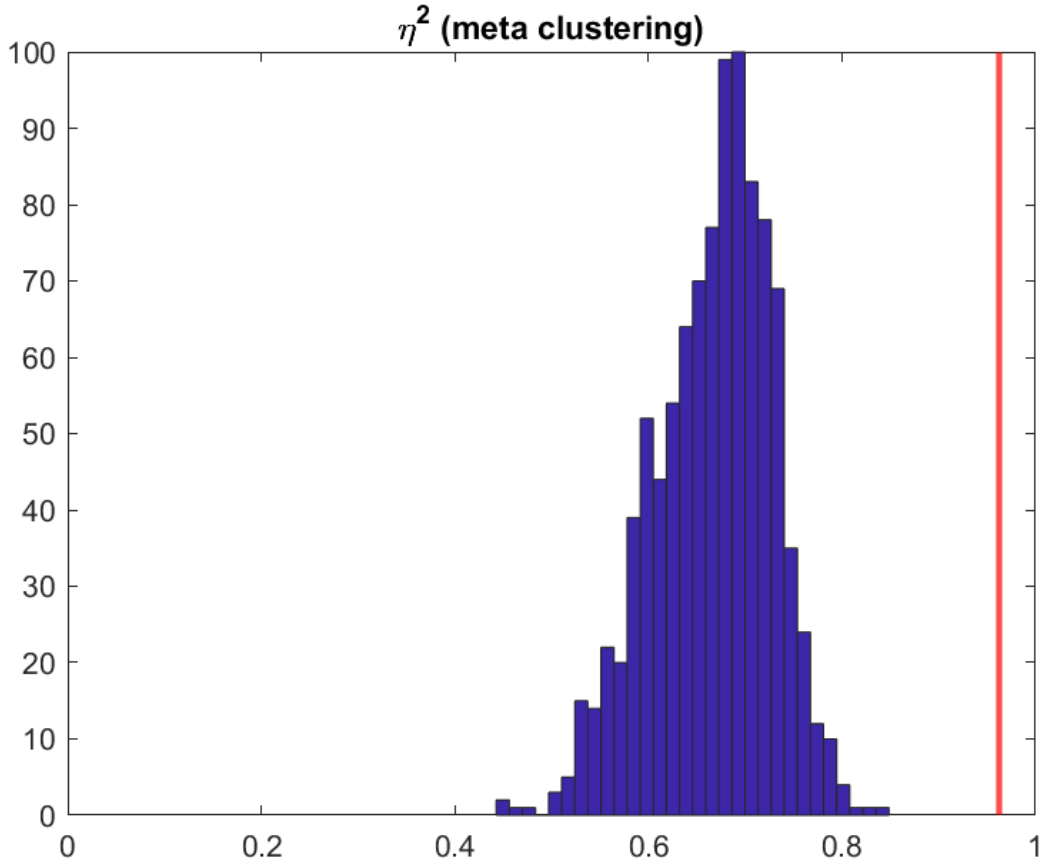


Sweden



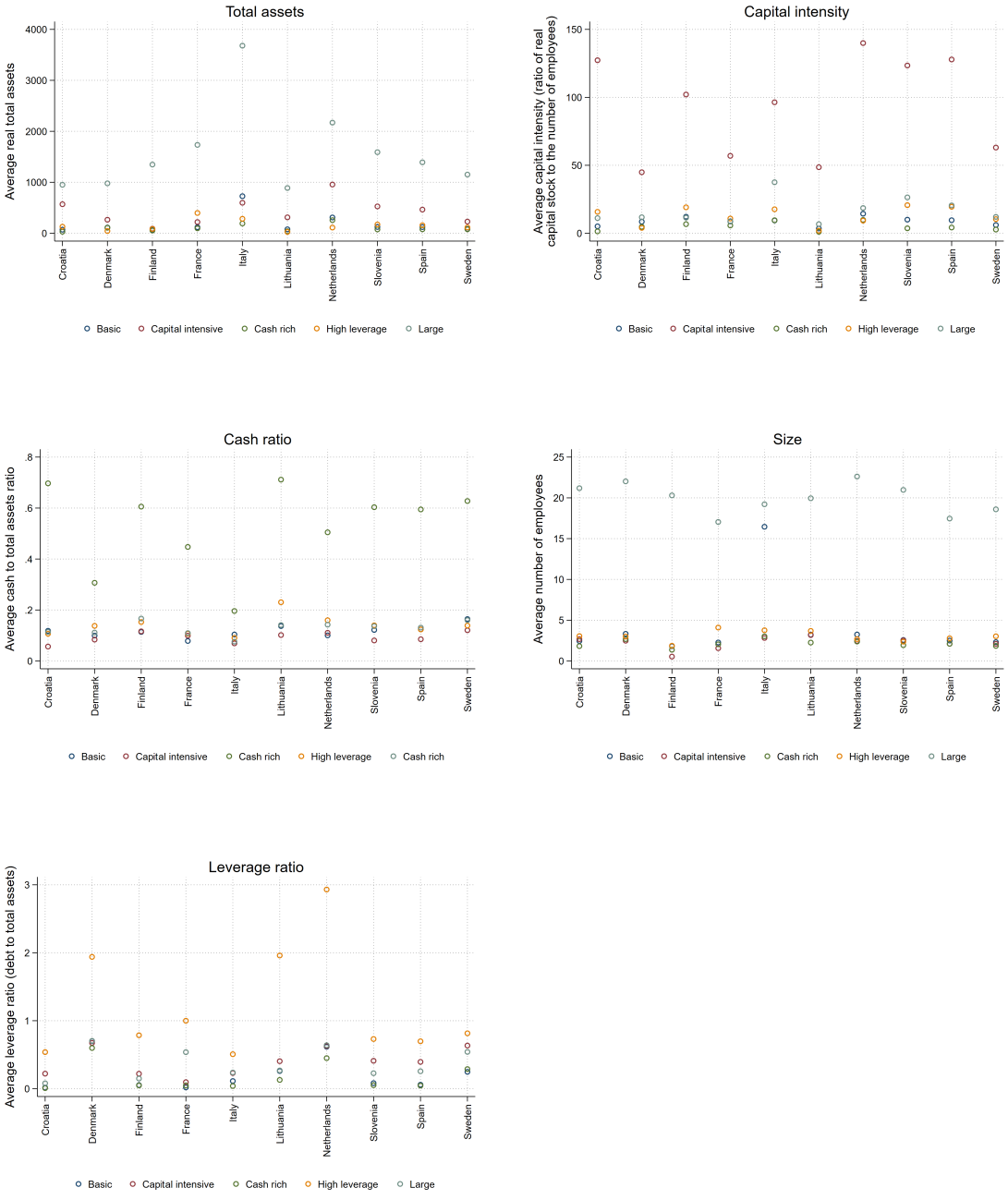
Notes: This figure shows scree plots resulting from the k-means cluster algorithm at the firm-level, for each country in the sample. On the x-axis, k indicates the number of clusters. The η^2 coefficient on the y-axis measures the proportional reduction of the within sum of squares for each cluster solution k compared with the total sum of squares.

Figure A6. Monte Carlo validation of the clustering procedure



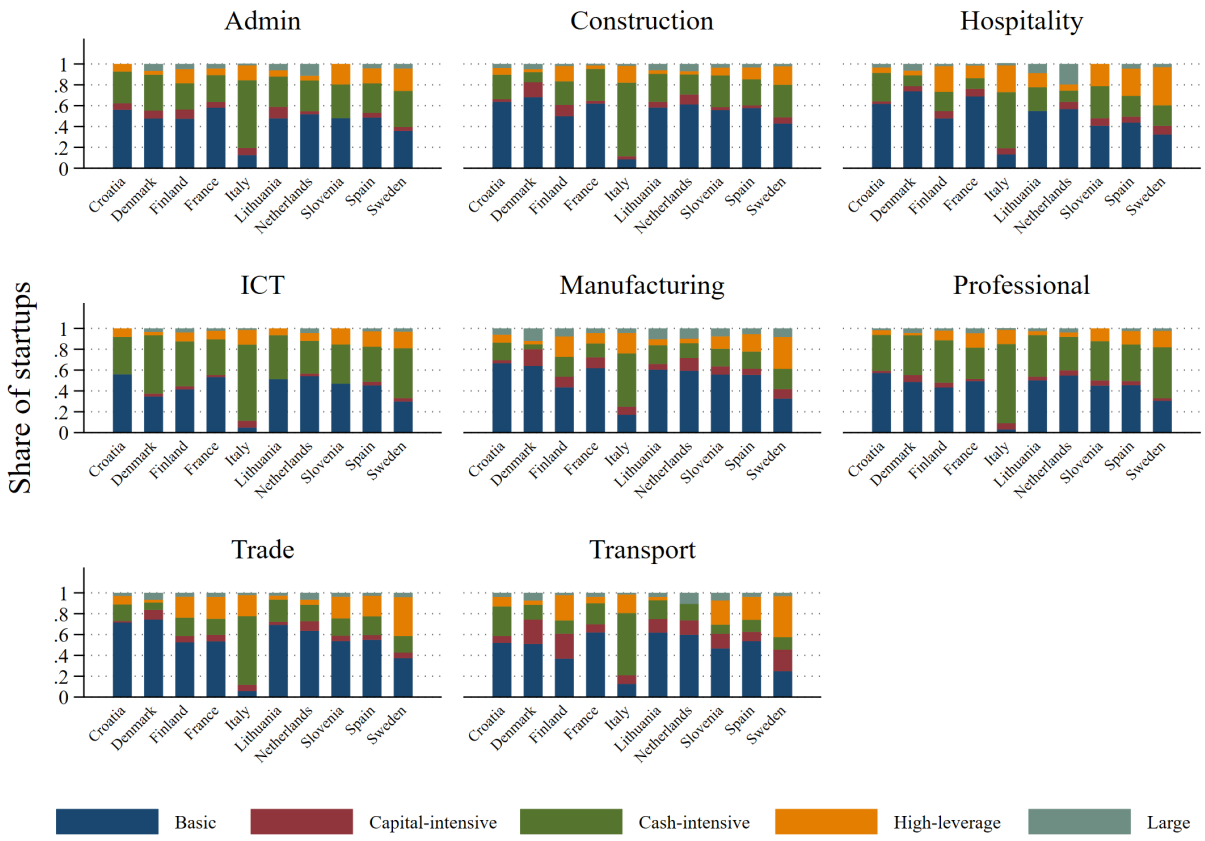
Notes: This histogram summarizes a Monte Carlo experiment consisting of 1,000 random draws for the cluster variables, for each country, with means and standard deviations as observed in the data by country. These draws are i.i.d. so that no clusters exist in the experimental data. Based on these artificially generated data, we then apply our clustering procedure as on the real-world data, i.e., first for each country separately, followed by a meta-clustering step across countries. Each time η^2 is computed (blue bars). The vertical red line indicates the true η^2 statistic based on the actual data. The experiment is based on six countries for which the standard deviations of all the cluster variables are available (Croatia, Italy, Lithuania, Spain, Slovenia, and Sweden).

Figure A7. Startup characteristics at the time of firm entry, by type and country



Notes: This figure presents the country-level mean of the five cluster variables for the five startup types in the year of firm entry.

Figure A8. Distribution of startup types by industry and by country



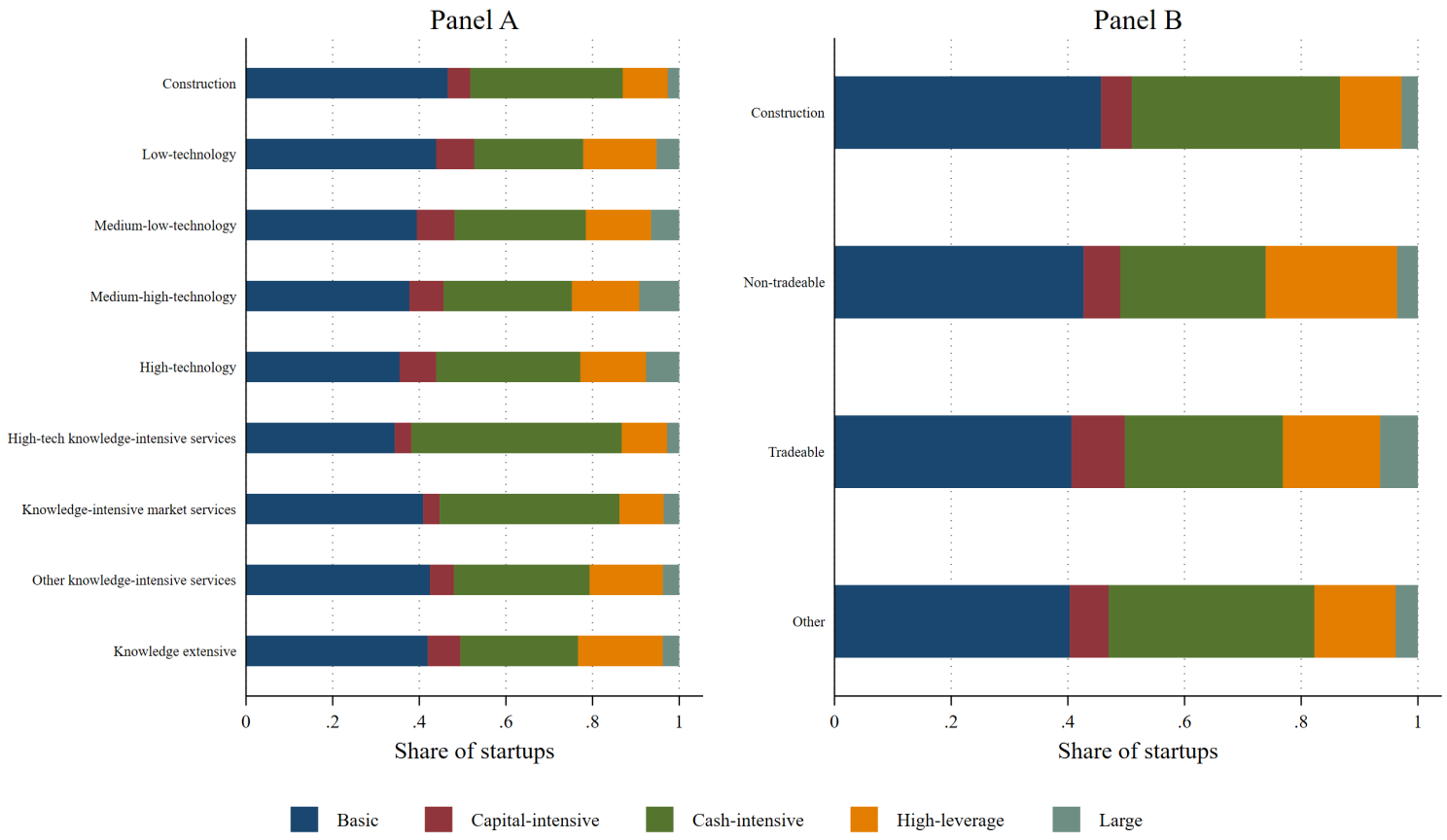
Notes: This figure presents the distribution of the startup population for individual one-digit NACE Rev.2 industries and countries across the five startup types. The startup population comprises startups that are just born (age 0) and shares are averaged over all cohorts available for each country.

Figure A9. Distribution of startup types by two-digit industry



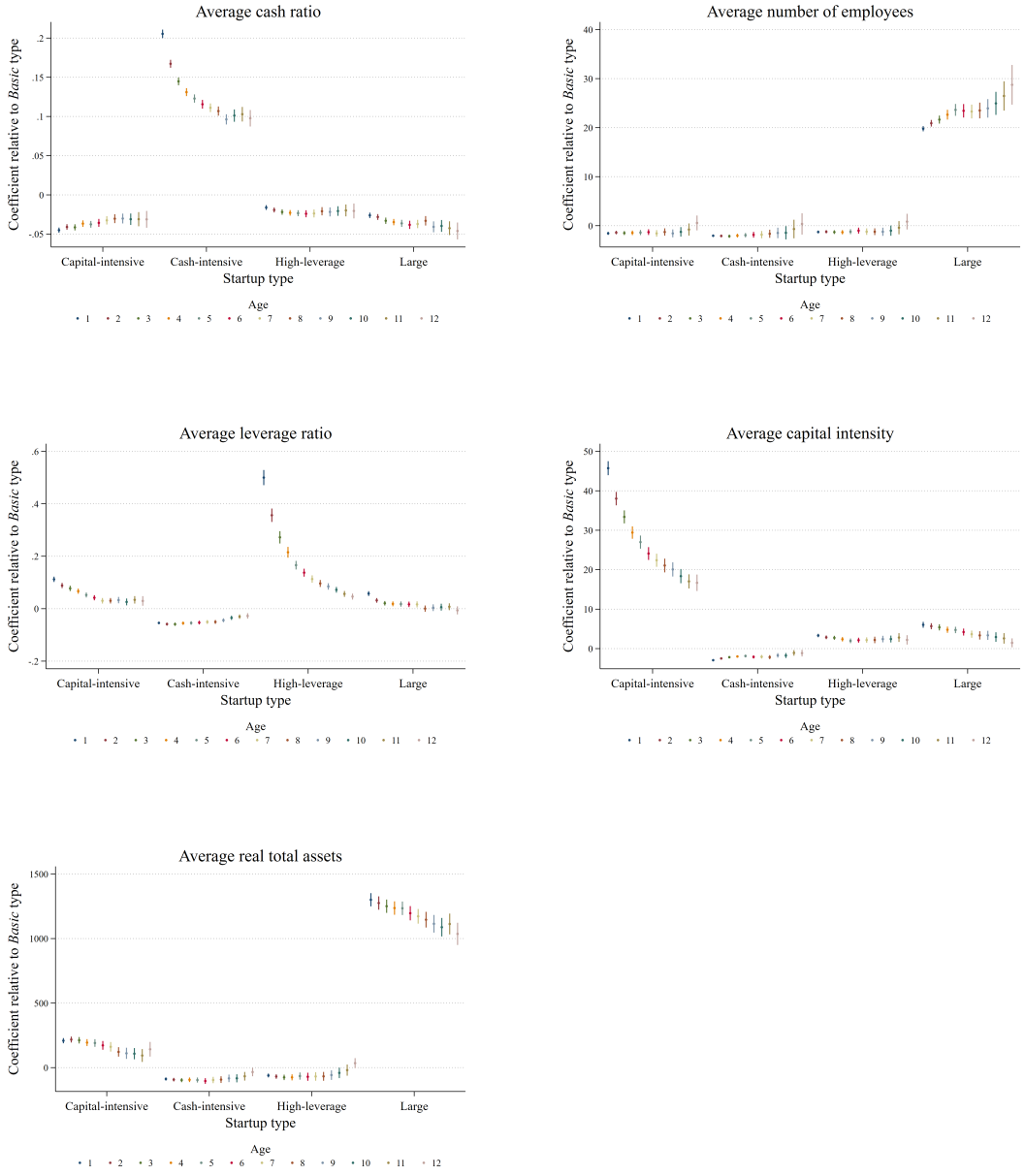
Notes: This figure illustrates the distribution of the startup population for two-digit NACE Rev.2 industries across the five startup types. The startup population comprises startups that are just born (age 0) in all cohorts available for each country.

Figure A10. Distribution of startup types by alternative industry classification



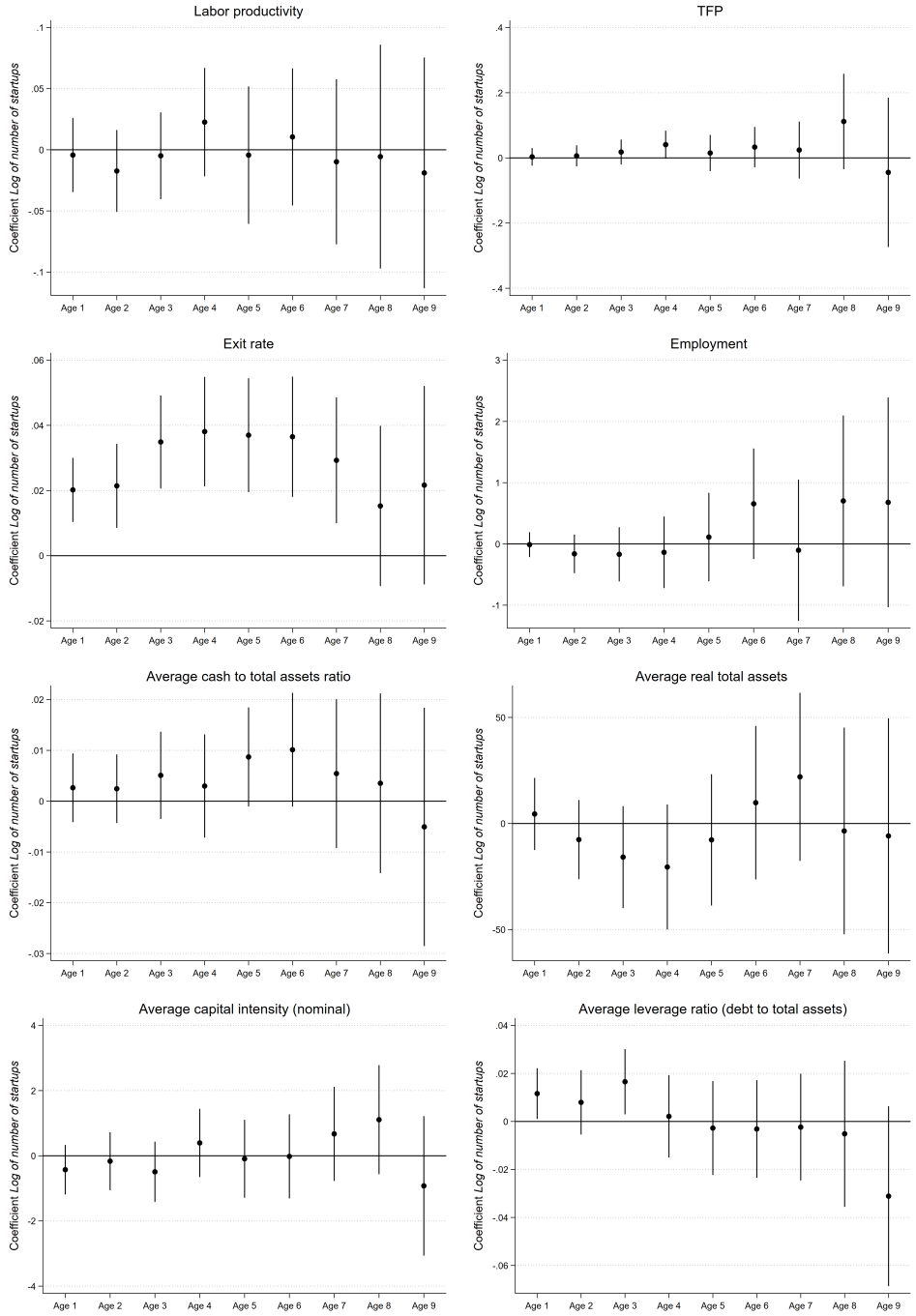
Notes: This figure illustrates the startup distribution for two different industry classifications. Panel A draws on Eurostat's classification of high-tech manufacturing industries and knowledge-intensive services. Using these classifications, we allocate two-digit industries into one of nine categories. Our data also includes 15 two-digit industries in construction, trade, transport, hospitality and administration that cannot be categorized this way. For the construction industries, we create a separate category while the remaining industries are grouped in a "knowledge extensive" category. Panel B draws on Mian and Sufi (2014) to assign two-digit industries to construction, tradable, non-tradable, and other. This is straightforward for 33 two-digit industries as all underlying four-digit industries have the same Mian and Sufi (2014) classification. For 22 other industries the underlying four-digit industries indicate at least two Mian and Sufi classifications. We then assign the categorization that is attached to the majority of four-digit industries within the more aggregated two-digit industry. The population comprises startups that are just born (age 0) in all cohorts available for each country.

Figure A11. The life cycle of startup types (one-digit industry level)



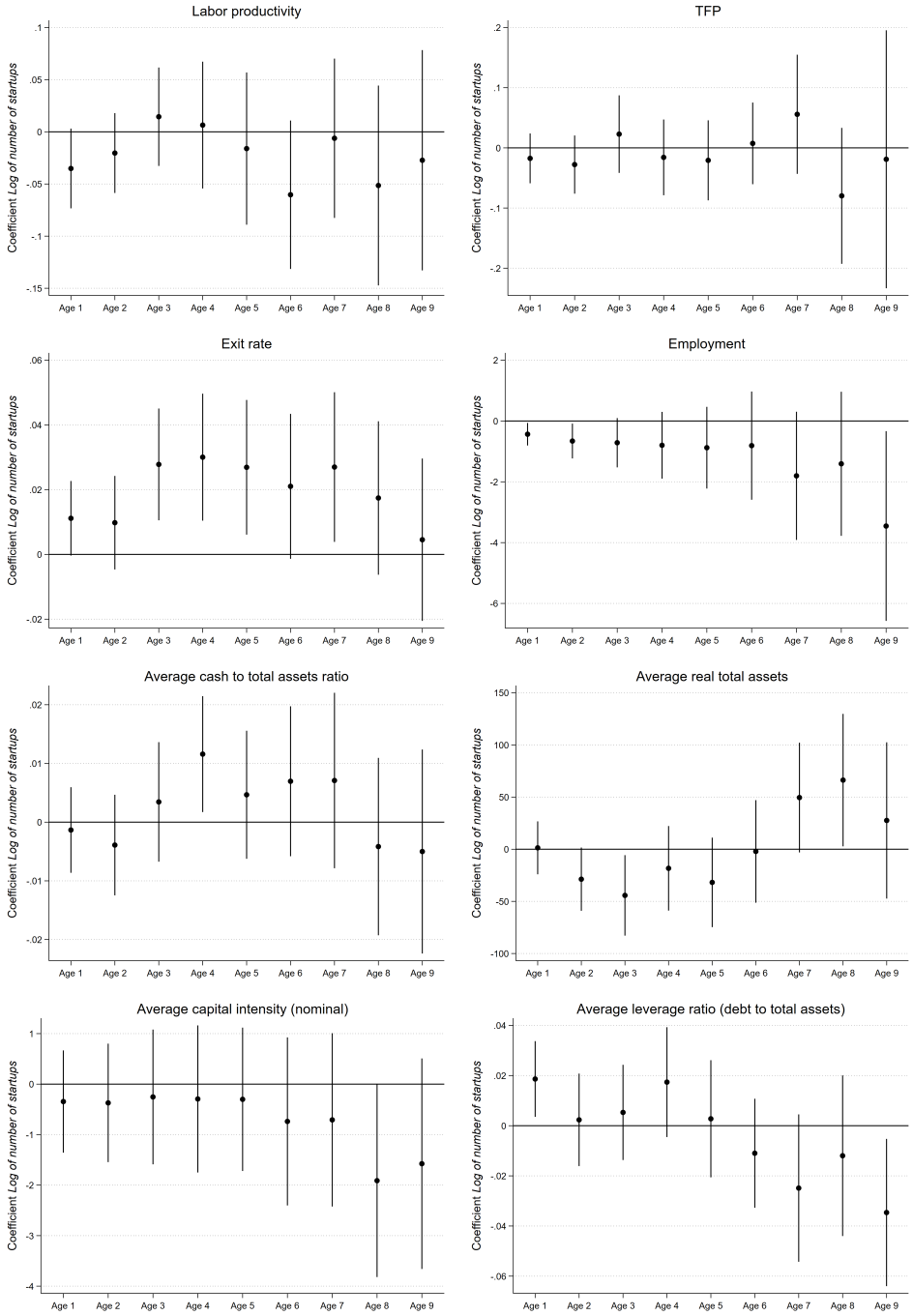
Notes: The panels in this figure summarize how the startup types develop during the first 12 years of their life in terms of the five clustering variables. Each panel corresponds to one clustering variable and plots the coefficients from 12 separate regressions where the dependent variable is this variable. Each regression is then run for an age group (age is 1, 2, ...12 years). For example, the first panel summarizes regressions in which the *Average cash to total assets* ratio is regressed on dummy variables for the startup types (the *Basic* type is omitted) as well as *Country* × *Cohort* and *Industry* fixed effects. The sample is the full panel dataset at the one-digit industry level. Standard errors are clustered by the *Industry* × *Country* × *Cohort* and whiskers indicate 95 percent confidence intervals. Capital intensity is measured in thousands of euros (PPP-adjusted) per employee; the cash and leverage ratios are expressed relative to total assets.

Figure A12. Entry selection: Cohort size and startup performance over time (2-digit sector)



Notes: This figure summarizes OLS regressions where the dependent variable is indicated in the figure titles. The independent variable of interest is the log number of startups in the cohort. *Exit rate* is the cumulative exit rate defined as 1 minus the ratio of number of startups that survived until t divided by the number of startups in *t-0*. Each dot corresponds to a partitioned regression for all startups of a particular age, as indicated on the x-axis, with interactive fixed effects for *Industry* $\times$ *Startup type* $\times$ *Country*; *Industry* $\times$ *Country* $\times$ *Cohort*; *Startup type* $\times$ *Country* $\times$ *Cohort*; and *Industry* $\times$ *Startup type* $\times$ *Cohort*. The sample is the full panel dataset partitioned by age and at the two-digit industry level. Standard errors are clustered by the *Industry* $\times$ *Country* $\times$ *Cohort* and whiskers indicate 95 percent confidence intervals.

Figure A13. Entry selection: Cohort size and startup performance over time (1-digit sector)



Notes: This figure summarizes OLS regressions where the dependent variable is indicated in the figure titles. The independent variable of interest is the log number of startups in the cohort. *Exit rate* is the cumulative exit rate defined as 1 minus the ratio of number of startups that survived until t divided by the number of startups in *t-0*. Each dot corresponds to a partitioned regression for all startups of a particular age, as indicated on the x-axis, with interactive fixed effects for *Industry* $\times$ *Startup type* $\times$ *Country*; *Industry* $\times$ *Country* $\times$ *Cohort*; *Startup type* $\times$ *Country* $\times$ *Cohort*; and *Industry* $\times$ *Startup type* $\times$ *Cohort*. The sample is the full panel dataset partitioned by age and at the one-digit industry level. Standard errors are clustered by the *Industry* $\times$ *Country* $\times$ *Cohort* and whiskers indicate 95 percent confidence intervals.

Table A1. Sample composition by country

Country	Sample period	Number of startups
Croatia	2003-2019	64,760
Denmark	2002-2018	114,195
Finland	2000-2019	126,554
France	2005-2007	210,033
Italy	2007-2018	322,893
Lithuania	2001-2017	47,322
Netherlands	2008-2018	63,729
Slovenia	2006-2019	23,435
Spain	2009-2018	236,192
Sweden	2004-2019	136,376
Total		*1,345,489*

Notes: This table shows the sample composition of the full panel of startups just born.

Table A2. Variable definitions

Variable	Definition
Capital intensity	Average ratio of real capital stock to the number of employees
Cash ratio	Average cash to total assets ratio
Employment	Average number of employees
Exit probability	1 minus the ratio of number of firms that survived until t divided by the number of startups at $t=0$
Labor productivity	Logarithm of employment-weighted average labor productivity, where firm-level labor productivity is defined as real value-added divided by number of employees
Leverage ratio	Average ratio of debt to total assets
Profit margin	Average ratio of operating profit (Earnings Before Interest and Tax (EBIT)) to revenue
Real total assets	Average real total assets (thousands of euros)
TFP	Employment-weighted average total factor productivity based on a GMM estimation following Ackerberg, Caves and Frazer (2015) and assuming a Cobb–Douglas production function
Wage per employee	Average wage per employee

Notes: Monetary variables are denominated in thousands of euros and PPP-adjusted. Real variables are deflated with sectoral price indices.

Table A3. Cluster means for alternative numbers of clusters

Startup type	(1) No. of employees	(2) Capital intensity	(3) Real total assets	(4) Cash ratio	(5) Leverage ratio
Panel A: Five clusters (baseline)					
Basic	4	8.56	166.59	0.12	0.23
Capital-intensive	2	93.18	405.12	0.09	0.41
Cash-intensive	2	4.67	92.46	0.54	0.18
High-leverage	3	12.90	122.97	0.14	1.18
Large	20	16.05	1488.30	0.13	0.34
Panel B: Six clusters					
Basic (small)	3	9.32	114.32	0.16	0.25
Basic (large)	18	8.51	611.74	0.14	0.32
Capital-intensive	2	94.05	375.65	0.09	0.42
Cash-intensive	2	4.10	89.61	0.61	0.18
High-leverage	3	12.69	120.39	0.15	1.22
Large	19	25.37	2225.78	0.12	0.34
Panel C: Seven clusters					
Basic (small)	2	5.66	88.86	0.35	0.23
Basic (medium)	3	14.98	176.95	0.09	0.28
Basic (large)	20	7.57	619.10	0.15	0.29
Capital-intensive	2	99.13	390.82	0.09	0.42
Cash-intensive	2	3.74	88.00	0.66	0.18
High-leverage	3	12.80	117.79	0.14	1.24
Large	17	28.19	2676.54	0.11	0.35

Notes: This table reports cross-country means of the five clustering variables for each cluster when the baseline clustering (with five clusters) is repeated with six and seven clusters. Means are unweighted and based on all available cohorts and countries. Real total assets are in thousands of euros, PPP-adjusted.

Internet Appendix B: Model details

To rationalize the existence of the different startup types, as well as their performance differences, we propose a model of heterogeneous firms. Specifically, we allow for heterogeneity in terms of (i) Total Factor Productivity, (ii) scalability with respect to capital and/or labor, and (iii) entry costs, fixed costs and initial assets. We consider the following policy variables:

- A corporate income tax (possibly differentiated by types): τ_π
- A labor tax: τ_l
- A capital tax: τ_k
- A loan tax: τ_b

All of these may be negative (subsidies).

1. Environment

The model period is one year and we denote firm age by a. We distinguish between three age categories: (i) potential entrants ($a = -1$), (ii) young firms ($0 \leq a < a^*$), (iii) old firms ($a \geq a^*$). Entrants and young firms are subject to real and financial rigidities, whereas old firms are unrestricted. Entrepreneurs have a discount factor $\Lambda \in (0, 1)$ and they can save outside of the firm in one-year bonds with a rate $r = \frac{1}{\Lambda} - 1$.

Technology. Let the production function of the firm be given by $y = A_a k^\alpha l^\gamma$, where TFP A_a varies positively with age (a) and is heterogeneous across firms. Moreover, α and γ are heterogeneous across firms (but constant over time). Capital depreciates at a rate $\delta \in (0, 1)$ per year. Production also requires the payment of a fixed cost c. Concretely, we assume that the fixed cost may be higher in the initial year, reflecting firm establishment costs. Subsequently, the fixed cost remains constant. TFP evolves as:

$$A_a = z + z_a$$

where z is a firm-specific but time-invariant component, and z_a is an age-dependent component which is constant across firms of the same age (and known to all).

Entrants. Starting a firm requires the payment of an entry cost θ, which is heterogeneous across entrepreneurs. Any firm starts with zero capital, and some initial level of assets m_{-1}.

A potential entrant observes its individual characteristics $\{z, \alpha, \gamma, \theta\}$ and decides whether to enter or not. If the firm enters, it must decide its capital level k which cannot be adjusted until becoming old (the firm pays an upkeep δk per period to compensate for depreciation). Upon entry, the firm can also take out a long-term loan with duration a^*, which is to be repaid in equal installments and comes with an interest rate $r_b(b) \geq \widetilde{r}$. We assume that $r_b(b)$ is weakly increasing in the loan amount b, and that $r_b(b) = \widetilde{r}$ when $b \leq 0$ (i.e. when the firm has positive savings). Here, $\widetilde{r} = \frac{1}{\sum_{a=0}^{a^*-1} \Lambda^{a+1}} - \frac{1}{a^*}$ is the "frictionless rate" for this kind of loan that is consistent with the subjective discount factor and the interest rate on savings outside the firm. In other words, whenever the firm borrows, it pays a premium that is increasing in the size of the loan and captures financial frictions. No further loans can be taken out and the firm cannot default on a loan or pre-pay it. Concretely we assume that $r_b(b) = \widetilde{r} + \tau_b + \zeta b$ for $b \geq 0$.

Young firms. Having made the initial decisions, the firm enters at age $a = 0$. As mentioned above, until age a^* the firm cannot re-optimize its capital and loan decisions. However, the firm can freely adjust its employment, l, in every period.

Old firms. At age $a = a^*$ a firm becomes fully grown. The firm's TFP level remains constant thereafter. Moreover, the firm becomes financially unconstrained (e.g., it can raise equity) and can freely adjust its capital stock.

2. Decision problems

We now formalize the firm's decision problems in the various phases. We will solve the model backwards and start with the fully-grown firms.

Old firms ($a \geq a^*$). The old firm's decision problem is equivalent to maximizing economic profits, which can be expressed as $\pi = (1 - \tau_\pi)\left(Ak^\alpha l^\gamma - (r + \delta + \tau_k)\, k - w(1 + \tau_l)l - c\right)$.

$$r + \delta + \tau_k = \alpha \frac{y}{k},$$
$$w\,(1 + \tau_l) = \gamma \frac{y}{l}.$$

We will denote the profits of a fully-grown firm by $\pi^*(z, \alpha, \gamma)$. Old firms exit exogenously at a rate ρ.

Young firms ($a = 0, 1, .., a^* - 1$). Now consider the firms in the "scale-up" phase, which lasts for a^* years. The firm does not exit during this phase. Recall that in this stage the firm can only adjust labor inputs, denoted l_a. It does so to statically maximize profits, given by:

$$\pi_a = (1 - \tau_\pi) \left(A_a k^\alpha l_a^\gamma - w(1 + \tau_l) l_a - (\delta + \tau_k) k - c \right) - \left(\frac{1}{a^*} + r_b(b) \right) b,$$

where $\frac{b}{a^*}$ captures repayment of the loan. The profit-maximizing choice of labor is again given by:

$$l_a = \left(\frac{w(1 + \tau_l)}{\gamma A_a} k^{-\alpha} \right)^{\frac{1}{\gamma - 1}}.$$

Taken together, the above two equations define a profit function for young firms, which we denote $\pi_a(A_a, k, b, w)$. The firm's short-term asset position evolves as:

$$m_a = (1 + r) m_{a-1} + \pi_a$$

where $a = 1, 2, .., a^*$ and initial assets satisfying $m_0 = m_{-1} + b - k + \pi_0$.

We define financial variables as follows. Let $\mathbb{I}_m$ be an indicator function such that $\mathbb{I}_m = 1$ if $m_a > 0$ (the firm has positive cash) and $\mathbb{I}_m = 0$ otherwise (the firm is in short-term debt). We measure cash as $cash_a = \mathbb{I}_m m_a$ and short-term debt as $debt_a^{ST} = -(1 - \mathbb{I}_m) m_a$. Moreover, long-term debt is given by $debt_a^{LT} = b_a(1 - \frac{a}{a^*})$. Accordingly, we measure the leverage ratio as $\frac{debt_a^{ST} + debt_a^{LT}}{k + cash_a}$ and the cash ratio as $\frac{cash_a}{k + cash_a}$.

Entrants. Immediately after entry, a firm decides on the capital stock and the loan, which cannot be re-optimized until age a^*. The value of an entrant firm is given by:

$$V(z, \alpha, \gamma, m_{-1}) = \max_{b,k,m_0} \sum_{a=0}^{a^*-1} \Lambda^a \pi_a(A_a, k, b) + \Lambda^{a^*} \left(\frac{\pi^*(z, \alpha, \gamma)}{1 - \Lambda(1 - \rho)} + k \right) + m_{-1} + b - k$$

subject to the following financing constraint:

$$k + m_0 \leq m_{-1} + b.$$

The constraint states that the firm must be able to finance the initial capital stock plus the initial fixed cost in production. The financing constraint is slack when $b \leq 0$.[40]

We can solve for the optimal choices from the first-order conditions, conditional upon the firm being in debt or not. A firm enters if and only if

$$V(z, \alpha, \gamma, m_{-1}) \geq m_{-1} + \theta.$$

We can now solve the problem as follows:

1. Solve for k from the following first-order condition, which holds conditional on the firm not being financially constrained ($b \leq 0$):

$$1 - \Lambda^{a^*} = \sum_{a=0}^{a^*-1} \Lambda^a \left((1 - \tau_\pi) \left(\alpha A_a k^{\alpha-1} l_a^\gamma - \delta - \tau_k \right) \right)$$

   where $l_a = \left(\frac{w(1+\tau_l)}{\gamma A_a} k^{-\alpha} \right)^{\frac{1}{\gamma-1}}$.

2. If for this solution $b = k - m_{-1} > 0$, then instead solve for k from the following first-order condition, which holds conditional on the firm being in debt ($b > 0$).

$$-\Lambda^{a^*} = \sum_{a=0}^{a^*-1} \Lambda^a \Big((1 - \tau_\pi) \left(\alpha A_a k^{\alpha-1} l_a^\gamma - \delta - \tau_k \right) - \left(\frac{1}{a^*} + r_b(k - m_{-1}) \right) - \frac{\partial r_b(k - m_{-1})}{\partial b} \cdot (k - m_{-1}) \Big),$$

   where again $l_a = \left(\frac{w(1+\tau_l)}{\gamma A_a} k^{-\alpha} \right)^{\frac{1}{\gamma-1}}$.

3. Compute $b = k - m_{-1}$ and compute profits at any age.

4. Compute the value of an entrant. The firm enters if and only if $V(z, \alpha, \gamma, m_{-1}) > m_{-1} + \theta$.

[40] In this case, the firm is at the margin indifferent between holding funds within or outside the firm, given that the interest rate is the same, and there is no implication for the firm's revenues when the firm is financially unconstrained.

3. Equilibrium

We now add a household sector to the model, allowing us to solve for the general equilibrium. The household has a subjective discount factor Λ, so that in the steady state $r = \frac{1}{\Lambda} - 1$ is given as above. We assume for tractability that households have GHH preferences over leisure and consumption. Optimal labor supply then takes the form:

$$l = \kappa_0 w^{\kappa_1}$$

where κ_1 is the labor supply elasticity, which can be calibrated externally, and κ_0 is a scaling parameter. The labor market clears if and only if:

$$l = \sum_j n_j \left(\sum_{a=0}^{a^*-1} l_{a,j} + \frac{l_{*,j}}{1-\rho} \right)$$

while the asset and capital market clear trivially. We consider budget neutral policy, imposing the following budget constraint:

$$\sum_j n_j \tau_{l,j} \left(\sum_{a=0}^{a^*-1} l_{a,j} + \frac{l_{*,j}}{1-\rho} \right) + \sum_j n_j \tau_{k,j} \left(\sum_{a=0}^{a^*-1} k_{a,j} + \frac{k_{*,j}}{1-\rho} \right) + \sum_j n_j \tau_{\pi,j} \left(\sum_{a=0}^{a^*-1} \pi_{a,j} + \frac{\pi_{*,j}}{1-\rho} \right) + a^* \sum_j n_j \tau_{b,j} \mathbb{I}_b \, b_j = 0,$$

where $\mathbb{I}_b$ is an indicator which equals one if and only if $b > 0$. In the policy exercises we jointly solve for w and one of the policy instruments to make the labor market clear and the government budget constraint hold. Asset markets clear trivially.